LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
Department "Institut für Informatik"
Lehr- und Forschungseinheit Medieninformatik
Prof. Dr. Heinrich Hußmann

Master Thesis

Design and Development of a Low-Cost Smartphone-Based Research Platform for Real-World Driving Studies

Matthias Christian Nefzger
Matthias.Nefzger@campus.lmu.de

Bearbeitungszeitraum: 15. 08. 2016 bis 19. 02. 2017
Betreuer: Dr. Ronald Schroeter (Queensland University of Technology)
Verantw. Hochschullehrer: Prof. Dr. Florian Alt

Abstract

Naturalistic driving studies as a method to explore driving and driver behaviour in a natural environment have become increasingly popular in road safety research. However, data acquisition systems needed for these studies are expensive and require a profound technical expertise for the installation. This thesis reports on an alternative approach towards more affordable naturalistic driving studies – a smartphone-based system called *Sensor Platform* that leverages the phone's sensors as well as external sensors to gather relevant driving data. In close cooperation with road safety experts, this project aimed to specify the requirements for such a system, develop a prototype, and evaluate from an user perspective as well as from a technical point of view. A focus group and an in-vehicle user study were conducted to gather the expert's feedback. In order to judge the accuracy of *Sensor Platform*, a comparison to an industry-grade data acquisition system was performed on the real road. The analysis of the study data suggests that road safety experts like the high usability and value the time savings. Yet, in comparison to industry-grade data acquisition systems, *Sensor Platform* is not on par when it comes to data accuracy, mainly due to simpler filtering algorithms. All in all, the thesis adds to the knowledge of mobile data acquisition systems while also providing a basis for future road safety applications such as real-time interventions.

Zusammenfassung

Sogenannte Naturalistic Driving Studies sind eine Methode um Fahr- und Fahrerverhalten in einer natürlichen Umgebung zu erforschen und sind in der Verkehrssicherheitsforschung zunehmend beliebt. Jedoch sind die nötigen Datenerfassungssysteme teuer und verlangen nach einer hohen technischen Expertise. Diese Arbeit präsentiert eine Alternative hin zu weniger teuren Studien: ein auf Smartphones basierendes System, *Sensor Platform* genannt, welches die Sensoren des Smartphones sowie externe Sensoren dazu nutzt, Fahrdaten zu sammeln. Dieses Projekt hat zum Ziel, in Zusammenarbeit mit Verkehrssicherheits-Experten die Anforderungen eines solchen Systems zu bestimmen, einen Prototypen zu entwickeln und das System aus der Nutzersicht sowie aus einer technischen Perspektive zu testen. Eine Fokusgruppe und eine Nutzerstudie im Fahrzeug wurden durchgeführt, um das Feedback der Experten zu sammeln. Durch einen Vergleich mit einem professionellen Datenerfassungsgerät auf mehreren Testfahrten wurde die Datengenauigkeit von *Sensor Platform* bewertet. Die Analyse der Studiendaten zeigt, dass die Experten die hohe Nutzerfreundlichkeit sowie die Zeitersparnis zu schätzen wissen. Allerdings ist *Sensor Platform* im Bezug auf die Datengenauigkeit nicht gleichwertig mit dem professionellen System. Der Hauptgrund dafür sind die einfachen Filteralgorithmen, welche in dem Projekt dieser Arbeit verwendet wurden. Die Ergebnisse der Arbeit tragen zum Wissen über mobile Datenerfassungssysteme bei und bilden die Grundlage für zukünftige Anwendungen im Bereich der Verkehrssicherheit, zum Beispiel Sicherheitsanzeigen während der Fahrt.

Acknowledgements

This thesis was conducted at the Urban Informatics Lab and the Center for Accident Research Queensland at Queensland University of Technology (QUT) in Brisbane, Australia. I would like to express my appreciation to Prof Florian Alt (LMU), Dr Ronald Schroeter (QUT), and Prof Marcus Foth (QUT) who made this stay happen.

Also, I want to thank everyone who supported me in my work at QUT in Brisbane, especially Ronny, Fabius, Patrick, Kai, and Anh.

This work was supported by a fellowship within the *FITweltweit* programme of the German Academic Exchange Service (DAAD).

Task Description

Since their introduction, naturalistic driving studies (NDS) have increasingly become a popular method of road safety experts. As opposed to driving simulator studies, NDS promise to yield natural data on driving behaviours that can help understanding the various causes for accidents. For example, the results of the first large-scale NDS showed that distraction plays a role in a surprisingly large number of events.

However, state-of-the-art data acquisition systems (DAS) for conducting NDS are expensive and often require a trained technician for the installation in vehicles. The task of this project is to explore an alternative, low-cost design of such a DAS. More specifically, the aim is to design and develop a mobile DAS in the form of a smartphone app that allows road safety researchers to conduct naturalistic driving studies on a budget and that is easy to use. The application should make use of the smartphone's internal sensors and cameras as well as potential low-cost external sensors to log critical driving events as a measure of driving behaviour. A modular architecture should allow future extensions of the application.

The project should include the completion of the following milestones:

- A review of current tools for logging driver behaviour data.
- Qualitative research (interviews and focus groups with road safety experts) and the translation of findings into requirement specifications.
- An iterative design and development cycle of the smartphone-based data acquisition system.
- A thorough evaluation of the system, both from a user-centric qualitative perspective with road safety experts (the prospective users of the system), and from a technical quantitative perspective to identify limitations around data accuracy.

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig angefertigt, alle Zitate als solche kenntlich gemacht sowie alle benutzten Quellen und Hilfsmittel angegeben habe.

München, 14. Februar 2017

....

Contents

1 Introduction

In 2013, 1.25 million people worldwide lost their lives in road traffic accidents, with millions more suffering serious injuries [1]. Even though notable progress has been made towards improving road safety legislation and vehicle safety, reports show that the rate of change is slow, especially in low-income countries [2]. In fact, 90% of road traffic deaths occur in low- and mid-income countries despite having just 54% of all vehicles. Even more, road traffic injuries are the leading cause of death among young people [1]. In light of these numbers, road safety experts aim to establish more effective interventions and legislation. For that purpose, researchers must have an accurate understanding of driver and driving behaviour worldwide.

One key research objective on the way to safer roads is to study the causes of crashes and near-crashes. However, over the last decades, research on crash causations has typically been based on police-reported accident statistics. These reports are in most cases derived from testimonies of those involved and are as such not objective representations of the accident and its causes. A major factor that is underrepresented in these self-reports is human error. Even though researchers concluded as early as 1977 that roughly 90% of crashes are caused by driver errors [3], finding the real causation of an accident involved costly and time-consuming on-scene investigations.

In the last decade, new approaches to data collection emerged, aiming to circumvent the drawbacks of self-reports. Besides the traditional road safety research methods of surveys, simulators, and observations, a novel method of data collection was established in the early 2000s [4]: Naturalistic driving studies are large-scale projects designed to collect vehicle sensor data over a long period of time. For that purpose, the participants' personal cars are equipped with a data acquisition system that unobtrusively gathers and stores driving data, usually over a timespan of several months. These systems generally include sensors to gather data on the movement and surroundings of the vehicle as well as cameras that record the driving scene from various angles. The resulting data offers insights into natural driving behaviour and footage of near-misses and crashes. For example, through the first large-scale naturalistic driving study, researchers were able to present evidence that in 93% of recorded crashes driver inattention was a contributing factor [4].

The main drawback of current naturalistic driving studies, however, is a very expensive study setup and a complex operation. For that reason, this type of study is usually conducted by multiple institutions who work together. In order to make naturalistic driving research more accessible, a number of recent research projects explored a more affordable method of naturalistic data collection: leveraging the ubiquity and low prices of today's consumer smartphones, researchers aim to collect driving sensor data through the phones' internal sensors (e.g. [5]). Mobile applications in this area have been developed pursuing different research aims, for instance real-time interventions [6], aggressive driving detection [7] or promoting an eco-friendly driving style [8]. Nevertheless, none of these projects aimed to develop a smartphone-based research platform that provides similar capabilities as expensive data acquisition systems. As a result, researchers looking for a low-cost study setup need to manually combine several applications and devices which creates overhead and poses compatibility issues.

1.1 Research Aims

The thesis addresses the described research gap and aims to explore a low-cost research platform for real-world driving studies. Through methods of Human-Computer Interaction, mainly user-centered design, the project includes the identification of requirements of road safety experts and the development of a system that matches these specifications. An evaluation of accuracy, reliability, and usability completes the iteration of the design cycle. The research questions explored in this thesis are as follows:

RQ 1: What requirements do road safety experts have for naturalistic data acquisition systems?

RQ 2: How can we design a smartphone-based data acquisition system (*Sensor Platform*) in a way that meets the requirements of road safety experts?

RQ 3 a): How do road safety experts judge the low-cost system (*Sensor Platform*) in terms of usability (i.e. effectiveness, efficiency, and satisfaction)?

RQ 3 b): How does the smartphone-based system (*Sensor Platform*) perform in terms of accuracy?

RQ 4: What are the benefits and disadvantages of a smartphone-based data acquisition system?

In summary, each of these questions is investigated during one of the four main phases of the project – requirements analysis, application development, evaluation, and analysis. The results of each phase directly feed into the process of the next phase.

The contribution of this thesis is twofold: In contrast to existing applications, this work specifically targets road safety researchers as the user group. Previous approaches towards smartphone-based sensing of driving data focused on single sensors or were designed to provide real-time feedback. No existing research project applied a user-centered approach in order to design a system that satisfies the requirements of road safety experts.

The second contribution comes in the form of the extensible data collection service. The system's code base will be made publicly available in order to facilitate future applications that rely on real-time driving data. Based on the service, new types of visualisations and interventions can be quickly prototyped and tested on the real road. Even more, existing concepts for road safety interventions such as *CoastMaster* by Steinberger et al. [9] can be studied outside of simulated environments. All in all, the two described contributions promise to add new insights to the field of road safety research and might benefit other researchers in the future.

1.2 Thesis Outline

The thesis is structured in the form of eight chapters, presenting the project in a roughly chronological order.

Following this introduction, the second chapter on **background knowledge** gives an overview of two topics that are highly relevant for this thesis, namely environment sensors and methods in road safety research. The section on sensors first defines important terms and concepts before detailing the available sensors in modern smartphones. Subsequently, road safety research methods along with several examples conclude the chapter on background knowledge.

The third chapter presents the **related work** in which this project is grounded. Firstly, the idea of sensor-based activity tracking is explained, followed by a detailed review of mobile phone based driving style analysis. In particular, the three kinds of data relevant to this field of research are described: vehicle data, environment data, and driver data. Lastly, several of the most important naturalistic driving studies are examined in regards to research aims and data acquisition systems.

Chapter four provides an overview of the **methodology** followed in this thesis. In summary, the approach utilizes methods from user-centered design and is based on the iterative interaction design model by Preece et al. [10]. A figure visualises the methodology and highlights where each of the four research questions is answered.

In the fifth chapter, a description of the **requirements specification** is given, starting with semi-structured expert interviews followed by a systematic analysis of the collected interview data.

Finally, the chapter specifies a comprehensive list of requirements on basis of related work and insights from the interviews. Moreover, expected limitations of the low-cost approach are identified and discussed.

The chapter about **application development** delves into the process of implementing a prototype version, including an Android application and a web server. Based on the requirements identified in chapter five, the system was designed by using paper sketches and UML modeling. Subsequently, the system's architecture is explained along with further details on the algorithms used. A list of features excluded from the implementation concludes the chapter.

Chapter seven reports on the **evaluation** phase and its findings. First, a focus group and a user study were conducted to gather feedback on the user interface and operation of the application. After that, the developed system is compared to an industry grade data acquisition system in multiple drives on the real road. Based on an analysis of the collected data, the chapter discusses strengths and drawbacks of the low-cost approach and provides a list of recommendations to follow in future iterations or related research.

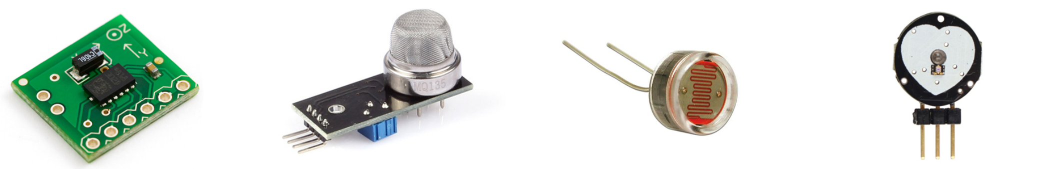
The last chapter **concludes** the thesis by critically reflecting on the developed platform, summarizing the research contribution and outlining possible directions for future work.

2 Background

This chapter provides an overview of the theoretical background for this thesis. Two main areas of background knowledge are relevant in this project: First, **environment sensors** and their capabilities are discussed, starting with a definition of terms and properties. With regard to the topic of this thesis, the available sensors in modern smartphones are presented. Finally, the chapter gives a brief overview of **road safety research** and the most common research methods in that field.

2.1 Environment Sensors

Today, sensors are used in countless areas, often in ways that few people are aware of. Applications include medical environments, transportation, communication, entertainment, manufacturing, and robotics (compare Figure 2.1). As this project heavily relies on sensor data to measure driving behaviour, this section clarifies the key terms and introduces the concept of sensor fusion – a method that is prominently featured in driving data acquisition systems.



(a) Accelerometer (b) Air Quality Sensor (c) Light Sensor (d) Heart Rate Sensor

Figure 2.1: Examples of environment sensors.

2.1.1 Definition and Terms

A sensor is defined as "a device that responds to a physical stimulus (as heat, light, sound, pressure, magnetism, or a particular motion) and transmits a resulting impulse (as for measurement or operating a control)" [11].

In most cases, sensors can be categorised as either analog or digital. Analog sensors are measuring a property of the real world continuously, outputting any value over the range of measurement. This value can then be transformed into a discrete digital signal in order to process it by a computer. In contrast, digital sensors represent physical attributes as a discrete signal. Figure 2.2 visualizes the difference between continuous and discrete signals. However, even in the age of digitalization, some purely analog sensors are still widely used, such as mercury thermometers.

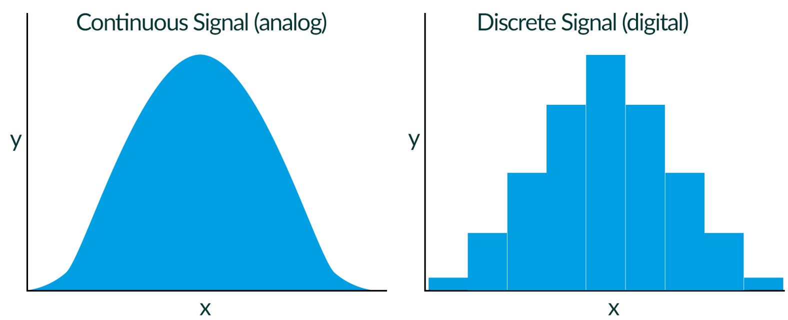


Figure 2.2: Continuous (analog) and discrete (digital) signals.

In most application areas, more than one type of sensor is used: **Sensor fusion**, sometimes also referred to as *data fusion* or *information fusion*, describes the process of combining sensory data from separate sensors to obtain information that is more complete, accurate or dependable than data from an individual source. Hal [12] provides a definition of *multisensor data fusion* as "the technology concerned with the combination of how to combine data from multiple (and possible diverse) sensors in order to make inferences about a physical event, activity, or situation".

A common example of sensor fusion is found in vehicular navigation systems guiding cars: location data from the Global Positioning System (GPS) is combined with inertial movement data from the vehicle and processed by an algorithm known as Kalman Filter [13]. Even more complex examples of sensor fusion with the Kalman Filter can be found in aerospace projects [14]. In these examples, sensor fusion improves the accuracy of information and allows predicting future values.

Sensors of the same type (e.g. accelerometers) often differ in their exact specifications; several properties can have an impact on the ability to sense environment data:

- The **value range** of a sensor describes the range over which it is functional and can report data. If the values exceed this range, the sensor is not able to capture the data anymore or reports wrong values.
- A sensor's **accuracy** describes how well it measures the surroundings in an absolute sense
- The **resolution** or **sensitivity** defines the smallest change a sensor is able to detect. The resolution may be of a finer granularity than the accuracy, resulting in a certain range of uncertainty.

Information about these three sensor properties is important in order to assess whether a specific sensor is capable of gathering suitable data.

Regardless of the exact sensor type, several forms of deviation or error can occur in sensor readings, as described by D'Amico and Di Natale [15]. The classical design of analog sensors has a major drawback. Due to unavoidable fluctuations in chemical, physical or biological properties, the analog signal is prone to **electrical noise**. Noise caused by environmental factors such as electromagnetic fields and radiation is the reason why the accuracy of a sensor can never be zero. Digital sensors minimize this error due to the instant conversion of a value to a discrete signal.

Another deviation to consider is the so-called **sensor drift**, which describes a slow and unpredictable low-frequency change over time. Sensor drift can, for example, be caused by aging electronic components. If the drift is known to be constant, the signal can be corrected to eliminate its influence on the data. An example for sensors that are often affected by sensor drift are the gyroscopes used in consumer electronics [16].

Finally, even though a sensor is almost always designed to be sensitive to only one specific environmental property, it often is sensitive to other physical, chemical, or biological factors as well. This behaviour is called **cross sensitivity**. It is therefore important to be aware of these other influences and take their effects into account. One approach to overcome cross sensitivity is to use sensor arrays.

These three examples of deviations show that when working with sensor data, the limiting factors must be considered in order to correctly interpret the resulting data.

2.1.2 Smartphone Sensors

The core technology used for this thesis' project are smartphones – an example for devices in our everyday lives that constantly make use of sensors. Modern phones are equipped with a variety of hardware sensors that are used to record the environment and react to changes such as light conditions, movement or even air pressure. Many of these sensors are potentially useful to describe driving behaviour. The following list provides a comprehensive overview of sensors commonly installed in recent smartphones.

- The **back-facing camera** is a high-resolution camera that allows taking pictures or videos. Recent smartphone models can often capture videos in up to 4k resolution and 240fps at a lower resolution.
- The **front-facing camera** is a wide-angle camera with a lower resolution compared to the back-facing camera. The lens is positioned directly next to the screen and is able to capture images or videos of the user looking at the screen. Exemplary applications using this camera are video-calling apps or user authentication methods that rely on facial features.
- A **Global Positioning System (GPS) sensor** allows retrieving the absolute device position up to a certain precision (approximately 5-10 meters [17]). The information about the position includes latitude, longitude, altitude and bearing.
- The **3-axis Accelerometer** measures the acceleration force applied to the device on three physical axes (x,y,z). The raw values include earth gravity.
- The **3-axis Gyroscope** measures the device's rate of rotation around each of the three physical axes (x,y,z).
- A **magnetic sensor** can be used to get measurements on the ambient geomagnetic field, including noise from an electronic device in the proximity.
- The **light sensor** is included to measure the ambient light level. Its most common use is to control the screen's brightness setting depending on the current light level.
- With the **proximity sensor**, the distance of a near object in relation to the device's screen can be estimated.
- A **microphone** to record sound.
- The typical sensors for connectivity are **WiFi**, **Bluetooth**, **Cellular**, and **NFC**.
- A **fingerprint sensor** for biometric authentication of the user.
- Some models also include special sensors to sense their environment such as a **thermometer**, a **hygrometer** or a **barometer**.

Based on these hardware sensors, additional virtual sensors can be derived through software sensor-fusion. Among the virtual sensors are:

- The earth's **gravity** on three axes. By combining values from the accelerometer and gyroscope, it is possible to obtain a precise measurement of these values.
- The **absolute device orientation** can be derived from a combination of gyroscope, accelerometer and magnetometer measurements where each of the hardware sensors is used to correct errors of the other sensors.
- The **linear acceleration** can be obtained by removing the effect of gravity from the accelerometer measurements. This can be done by applying a low-pass filter.

2.2 Road Safety Research

Next to environmental sensors, existing road safety research methods form the second area of important background knowledge. In order to design a successful data acquisition system for a naturalistic environment, it is important to understand the general advantages and limitations of the approach. This can be done by comparing it to other methods in the field.

The history of road safety research dates back more than 90 years. One of the first recorded studies on road safety was carried out in 1929 [18], investigating individuals' proneness to be involved in accidents. Since then, many different methods were developed to examine various aspects of road safety. In the following, three major methodological groups, namely **interviews**, **simulator**

studies, and **real-world driving studies**, will be discussed along with specific examples. The presented methods are ordered in regard to their relation between involved costs and applicability to the real world, as can be seen in Figure 2.3.

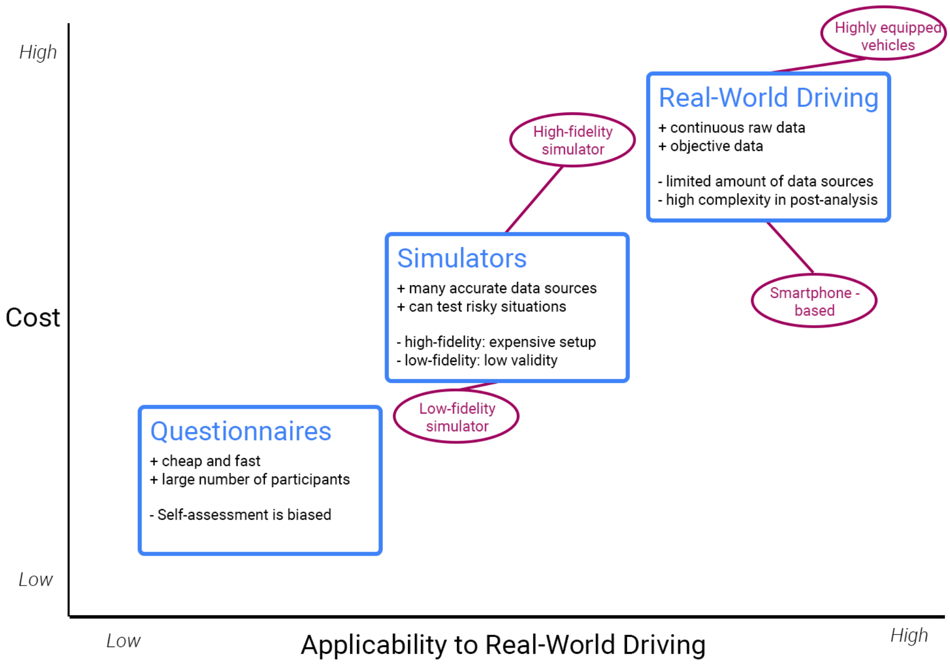


Figure 2.3: Comparison of road safety research methods in terms of cost and applicability, adapted from [7]

2.2.1 Interviews and Questionnaires

Out of the presented methods, interviews or questionnaires have the lowest costs. Questionnaires offer an inexpensive way to gather data on drivers' personalities and driving styles to estimate their risk of having an accident. A clear benefit compared to other methods is the large number and range of participants that can be reached. Over the last decades, questionnaires have been developed that specifically target different aspects of road safety. Popular examples include the *Driving Style Questionnaire (DSQ)* [19], the *Decision-Making Questionnaire (DMQ)* [20], the *Manchester Driver Behaviour Questionnaire (DBQ)* [21], and the *Driver Attitude Questionnaire (DAQ)* [22].

For example, the DBQ aims to predict future accident involvement by asking participants about their typical behaviour. The questionnaire has been shown to be a significant predictor of self-reported accidents [23]. However, further findings suggest that the DBQ only accurately predicts self-reported incidents, not crashes in general [24]. In addition, the methods to analyse the answers vary substantially between different studies, making it hard to compare the reported findings. As a result, af Wåhlberg et al. [24] conclude that the DBQ may not be as successful in predicting crashes as often claimed.

Due to the nature of this data collection method, interviews and questionnaires are not suitable to collect real-time driving data. In most cases, questionnaires and interviews require the participant to self-assess the driving which can lead to biased data. An analysis by West et al. [25] showed that while self-reports on speed are mostly reliable, self-judgements of driving aggressiveness differed significantly from the judgement of an expert observer. Another drawback of questionnaires that try to explore specific events or feelings during a drive is the effect of selective memory. Participants have to recall the situation after the drive and are likely to forget or add certain aspects to their report (as explored by Thomas and Diener [26]). In summary, questionnaires in road safety allow for relatively affordable and large-scale studies. However, it is important to account for certain drawbacks such as the bias in self-assessments.

2.2.2 Simulator Studies

A different approach take driving simulator studies, as they are designed to provide a driving environment similar to the real world without exposing participants to the obvious risks. The degree of similarity between the simulator and the real-world driving environment is expressed as fidelity. The fidelity of simulator setups ranges from laboratory-based low-fidelity simulators to expensive and complex high-fidelity simulators.

Low-fidelity driving simulators are based on a low-cost hardware setup, often featuring only a single screen displaying the driving simulation and a steering wheel connected to the computer. Depending on the specific study, extensions such as pedals or additional displays can be easily installed.

In contrast, **high-fidelity driving simulators** are often constructed around a model of a real car placed on a motion platform. With this approach, driving feedback such as acceleration and road curvature can be simulated by moving the simulator platform. These expensive setups usually feature a wide, often projection-based curved display in front of the driver to simulate peripheral vision. Figure 2.4 shows an example of a high-fidelity driving simulator.

Independent of the fidelity, there are many **reasons to use driving simulators** in automotive research: One key benefit of simulated environments is the ability to test driving situations that are too difficult or dangerous to test on real roads. As such, simulator studies are often conducted to test hypotheses around psychological research questions such as the effects and causes of risky driving behaviours. Another use case of simulated environments is the evaluation of planned road designs such as tunnels where it is not yet possible to test the conditions on a real road. In addition, simulations allow examining situations that seldom occur in the real world by staging the events leading to the rare situation.

Figure 2.4: High fidelity driving simulators: (a) A real car in front of a curved projection, (b) A motion platform simulates accelerations.

However, studies conducted in driving simulators, even high-fidelity ones, always bear the risk of **misrepresenting reality**. The validity of simulator study results refers to their applicability in a real-world driving setting. To assess the validity of results it is important to differentiate between absolute and relative validity [27]. Results have a high absolute validity if the absolute size of the investigated effects is similar to real world values. Relative validity describes the correlation between normalized values and general directions.

Various research studies explored the **validity of simulator studies** by comparing results to studies in the real world: Hallvig et al. [28] compared a simulator setup to real road driving investigating driver sleepiness. They conclude that while the absolute validity of the simulator results was low, relative validity was at a good level. In the case of this comparison study, the response patterns to night driving were similar even though the absolute values differed. After a survey of comparison studies, Kaptein et al. [27] draw a similar conclusion: For a mid-level, fixed-based driving simulator the results for behavioural variables were generally valid in terms of relative validity but absolute values were not transferable to the real world. However, some aspects such as route choice decisions were shown to be the same in real traffic and the simulation. For physiological measurements, Reimer and Mehler [29] report that the findings meet the statistic criteria for both relative and absolute validity.

2.2.3 Real-World Driving Studies

Real-world driving studies promise to generate the most accurate and unbiased data. Two substantially different data collection methods for real world driving data exist:

The first one, **expert-based examination**, requires a researcher to be present in the car to observe both driver and driving behaviour. The researcher can make notes on driving situations, give navigation directions and control the experiment. Additionally, this method allows gathering insights into driver situation awareness through verbal protocol analysis where drivers are encouraged to “think out loud”. A drawback of this data collection method is that drivers are potentially distracted by the researcher’s presence in the car. Furthermore, participants might change their driving behaviour when they are aware of the observation. Thus, expert-based examination offers good experimental control and does not require any special equipment but may change the subject’s driving behaviour.

The second approach to measuring real-world driving data is based on various **sensors installed in the car** or worn by the driver. This approach can be further distinguished by how obtrusively the sensors are placed. Some real-world driving studies require participants to drive in a car that is modified in an obvious way. This includes, for example, invasive methods to measure physiological data such as in [30]. In those cases, the sensor technology usually is too big to hide or requires to be directly fixed to the human body. Of course, the driver’s constant awareness of the technology during the study must be taken into account when drawing conclusions.

In contrast, unobtrusive methods build on the assumption that the driver forgets about the technology in the car. For so-called **Naturalistic Driving Studies (NDS)** the participant’s car is equipped with various small sensors that collect large amounts of real-time data. The modifications to the car are designed to be as non-invasive as possible. Data acquisition systems (DAS) of NDS often consists of several cameras, inertial measurement units (IMU), radar and a central computing device (see Figure 2.5 for an example). After installation, the DAS remains in the vehicles for a longer period of time, usually several months or even years. During the study, the system stores the data on a hard drive or sends it to a remote server for analysis. Due to the complex hardware setup, the long period of data collection and the regular need of maintenance, NDS are very expensive to conduct. Consequently, naturalistic driving studies are usually funded by national governments and executed by a cooperation of research institutes. In some cases, the raw data was made publicly available after the initial analysis, enabling other research groups to explore the data set (e.g. the 100-car NDS [4]). In recent years, projects aiming to develop mobile data collection systems try to overcome the high costs and complexity (e.g. [5], [6]).

Naturalistic Driving Studies differ from other data acquisition methods in road safety as there is no experimental manipulation at all. According to Valero-Mola et al. [31], the presence of the DAS is likely to have little or no influence on driving behaviour. Although this assumption was extrapolated from study results in related fields, it can be assumed that naturalistic driving data provides researchers with the most natural and unaffected data sets. Of course, NDS are mostly limited to data values that can be automatically collected through sensors even though some studies included methods for self-reports after special events. This implicates that NDS are very well suited to study events and habits in daily driving but are not designed to evaluate new technology or interfaces.

Figure 2.5: ERGONEERS vehicle testing kit for naturalistic driving studies.

The main goal of this chapter was to provide an introduction to the theoretical background of this thesis: environment sensors and road safety research methods. **Environmental sensors** are a crucial part of automation. Without the ability to acquire data about the environment, computer systems would not be able to perform informed calculations. Moreover, sensor data can build the basis for research in many different disciplines. Road safety and health research are two prominent examples. The section also revealed that it is important to be aware of existing limitations of sensors such as noise and inaccuracy. Consecutively, three different research **methods in road safety** were introduced: questionnaires, driving simulators, and real-world observations. A comparison in regard to costs, applicability and complexity revealed major differences. However, all methods, if applied correctly, can help road safety researchers to gain informative insights. It is important to notice that the approaches should not be thought of as mutually exclusive but as data sources that complement each other.

3 Related Work

In order to understand the current state of research on activity tracking and driving data, this chapter presents related projects and applications in three areas: the first section focuses on the **quantified self movement**, a form of sensor-based activity tracking with the aim to improve well-being. It is followed by a section about findings on **mobile driving style analysis**, based on the categories of vehicle, environment and driver data. The third part presents previous large-scale **naturalistic driving studies** and highlights differences in the study setups and data collection mechanisms.

3.1 Quantified Self Movement

As discussed in Section 2.1.1, current models of smartphones feature a variety of hardware and virtual sensors. Besides the obvious use cases for these sensors such as detecting screen orientation changes or adapting the display brightness, developers and manufacturers have come up with a broad range of applications that build on sensor data. One such example is the **Quantified self movement**, also known as "lifelogging": consumer technology such as mobile phones and wearables are used to gather data on various aspects of an individual's life, usually with the aim of self-improvement. A large part of this data is automatically collected by sensors, complemented by manual user inputs, e.g. details on food consumption. This form of sensor-based activity tracking is a close relative to measuring driving behaviour on smartphones; it is simply an example of how the same types of sensors are used in different areas. The raw data from the individual sensors becomes meaningful only by fusing several sources to create a bigger picture.

The two main areas of data collection in the Quantified self are *fitness* and *health*. Quantified self applications allow tracking aspects of daily life such as physical activity, food consumption, sleep patterns, heart rate, and even subjective data such as the daily mood [32]. Some developed systems aim for an even more complete picture and incorporate medical lab results or medication tracking. The common goal of these applications is to identify areas for possible improvement in order to increase fitness, health, and happiness. Users of activity tracking applications named "improving health", "improving work efficiency and cognitive performance", and "finding new life experiences" as their main motivations [32]. As such, the whole quantified self movement can be seen as a part of a new patient-driven healthcare model that aims to introduce a higher level of information flow, transparency, customisation, and responsibility-taking to the traditional healthcare system [33]. Hence, fitness and health are the core areas of data collection in order to provide users with enough information to take more responsibility and make informed decisions.

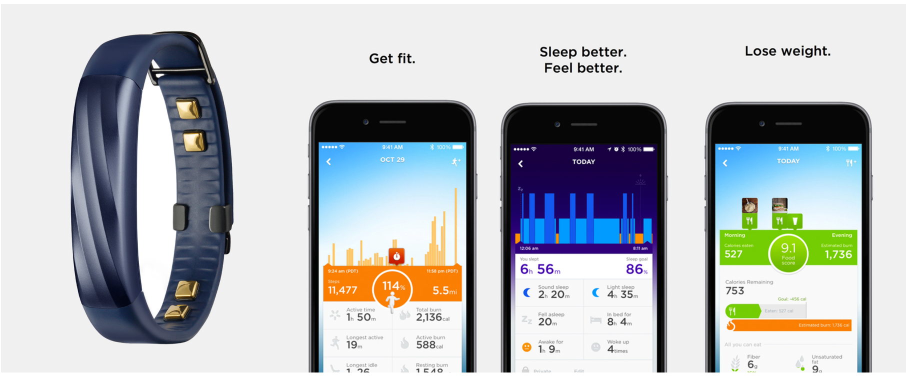


Figure 3.1: Fitness wearable Jawbone UP3[1] and its corresponding smartphone app.

[1] https://jawbone.com/store/buy/up3, accessed 30.08.2016

Self-monitoring as the process of recording one's own behaviours, thoughts, and feelings is not new in itself. As a part of behavioural psychology, first research projects were done as early as 1970 [34], however, these projects were always set in a clinical or research environment. The sensors in today's consumer devices allow any user to gather data that they were previously unable to access. This change is further supported by a whole new category of devices. So-called *fitness wearables* are unobtrusive computing devices designed to be worn on the body throughout the whole day and even at night. Popular examples for such trackers are *FitBit Blaze*[2], *Garmin Vivosmart*[3], and *Jawbone UP3* (see Figure 3.1). These devices usually include an accelerometer, a heart rate sensor and in some cases an additional galvanic skin response (GSR) sensor. Notably, this type of data is also becoming increasingly important when it comes to understanding driver behaviour (e.g. in [35]). Therefore, fitness wearables might be a valuable addition to driving data acquisition systems. In general, modern technology has made it possible for consumers to collect data that has previously been reserved to medical or research environments.

Similar to other data collection methods, the *process underlying the quantified self movement* can be divided into several stages. Swan et al. [36] developed a four-step model, consisting of data acquisition, software processing, information visualisation, and action-taking. Similarly, a five-step process was defined by Li et al. [37] (compare Figure 3.2). The first stage is entered with an individual's motivation to collect personal information and the informed decision on how to collect it. Based on the second stage, i.e. the raw data acquisition through sensors, software algorithms process and fuse the data to create more meaningful information. This information is then presented to the user in a graphical user interface, often including graphs and figures, which allow the user to reflect on the gathered information. The last step of the process is described as the "action stage". At this point, the individual can make a decision based on the visualized data and the advice given by the system. This multi-stage concept is applied by most popular applications for self-tracking.

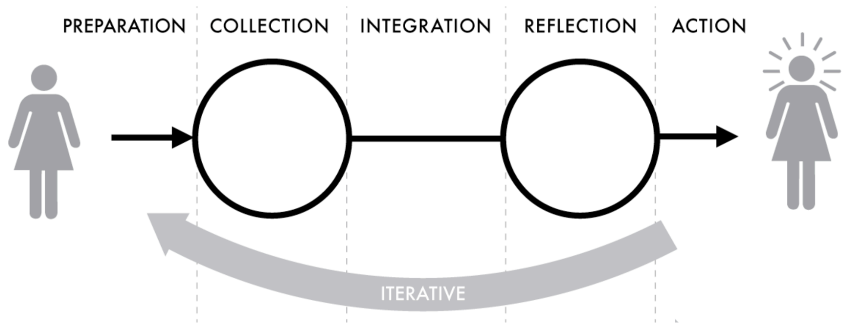


Figure 3.2: The five stages of quantified self applications, adapted from [37]. The big circles indicate the two stages that profit the most from system-driven processes that reduce the effort a user has to put in.

Exemplary applications highlighting the quantified self process were developed to track different areas of everyday life: *Readit* is a system that can track eye movements while reading to record and analyse reading habits [38]. The authors conclude that their methodology could be applied to a wide variety of cognitive tasks, supporting users in learning, concentrating, and retaining information. Another example is *Sense2Health*, an application that visualizes daily exposure to environmental pollution such as noise [39]. The method of sensor-based activity tracking has also been identified as a potential source of information to detect movement disorders and to assess surgical outcomes [40].

[2]https://www.fitbit.com/blaze, accessed 30.08.2016

[3]https://buy.garmin.com/en-US/US/sports-recreation/health-fitness/vivosmart-hr-/prod548743.html, accessed 30.08.2016

To summarise, the quantified self movement leverages sensors in consumer electronic devices such as phones and wearables to track an individual's daily activities. The main motivation for users is the promise of increased fitness, health, and wellbeing through the identification of areas for improvement highlighted in the sensor data. In the scope of this thesis, the quantified self movement shows that long-term data collection through smartphone sensors is a valuable source of information. The datasets allow classifying an individual's behaviour in regard to certain aspects through data fusion, a method that is very similar to naturalistic driving studies. Even more, the phone sensors used to track fitness and health are in many cases the very same that can be leveraged to measure driving style.

3.2 Driving Style Analysis

Road safety researchers have always been interested in the way people drive vehicles and how driving styles differ under varying conditions [41]. Ongoing research is conducted to understand the causes and effects of a driver's style of controlling a vehicle. Besides traditional methods such as the *Driving Style Questionnaire* (see 2.2.1), many research projects in recent years have been focused on the **automatic extraction of driving style and behaviour** based on sensor data. As an alternative to dedicated data acquisition systems, smartphones have become a popular platform. Instead of installing additional sensors and computers in the car, researchers leverage the ubiquity of smartphones in our everyday lives to gather driving data through the phone's sensors. As such, the approach of measuring the driving style through sensors in mobile devices is similar to the quantified self as discussed in the previous section. More specifically, some of the available applications that analyse driving behaviour are targeted at individual drivers, promoting to improve the driving style towards safer [42] or more eco-friendly habits [8].

Other examples of projects in this field of research include providing real-time feedback on drowsiness [6], measuring the driver's aggressiveness [7], and increasing safety at intersections [43]. Moreover, even though they all use a smartphone as the main DAS, the developed applications differ substantially in regard to which sensors are used for data collection. Some projects collect only inertial sensor data while others fuse a number of inputs such as video streams and data gathered from additional devices.

In general, automatically collected driving data can be divided into three main categories:

1. **Vehicle data** (e.g. vehicle movement, location, direction, speed)
2. **Environment data** (e.g. traffic infrastructure, following distance, lane keeping)
3. **Driver data** (e.g. distraction, drowsiness, arousal)

This classification is applicable not only to smartphone-based systems alone but to driving data acquisitions in general. However, in light of the research objectives of the thesis, the main focus of this section is on projects exploring low-cost approaches. The following subsections provide a definition of these categories and present previous academic work as well as existing applications or products.

3.2.1 Vehicle Data

Vehicle data includes all measurements that describe the vehicle's movement such as location, speed, direction of movement, longitudinal and lateral accelerations and turns. Additionally, vehicle data encompasses information on the vehicle's state of operation, in particular engine usage, fuel consumption, pedal use and steering wheel angle.

The most important data source for sensing vehicle data are the smartphone's inertial sensors. As presented in Section 2.1.1, modern phones come with a built-in accelerometer, gyroscope, and magnetometer. By fusing the readings from these sensors, it is possible to understand the vehicle's movement.

Accelerometer data allows monitoring the values of *longitudinal* and *lateral* acceleration. Figure 3.3 shows a visualisation of a vehicle's coordinate system. The longitudinal readings describe the acceleration in the direction of movement (the forward or backward direction) and can be processed to extract critical driving events such as hard accelerations and braking. Previous implementations frequently used thresholds (e.g. [43]) or time-based methods (e.g. [5]) for this processing step. In some cases, researchers established several levels for driving events, e.g. ranging from *low* to *high* [42], thus adding more context to the data.

The lateral acceleration of a vehicle is the change in sideway velocity and is a potential indicator for turns. However, to accurately detect turns, accelerometer data alone is not sufficient. A sensor fusion approach that combines the accelerometer data with the smartphone's rotation must be applied. Rotation data can be calculated based on the gyroscope and accelerometer readings. In summary, accelerometer data is an essential part of all systems monitoring the vehicle movement.

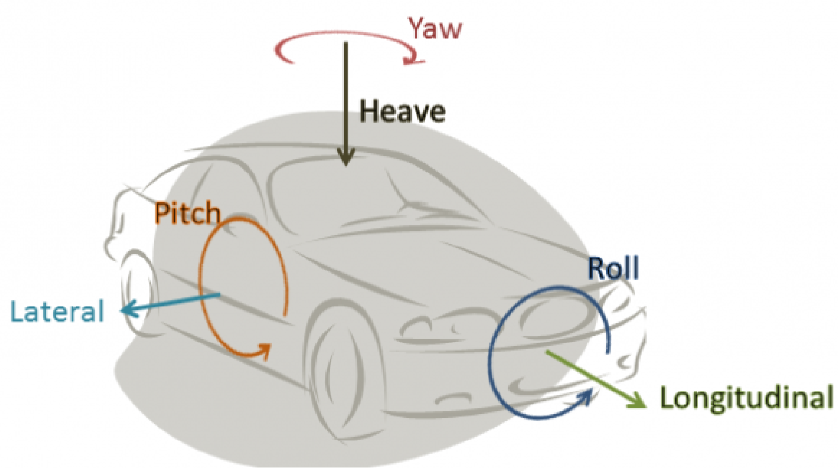


Figure 3.3: A vehicle's movement and rotation axes.

The **vehicle's speed** is the second important information about vehicle movement and can not be extracted from the accelerometer. Instead, the speed can be estimated through a time-series of GPS signals. The frequency of GPS readings on smartphones is usually much lower compared to other hardware sensors (around 1Hz), hence requiring further processing and extrapolation. A downside of using GPS to determine the vehicle's location and speed is the varying accuracy of the data. Zandbergen [17] found that consumer-grade GPS sensors installed in smartphones have a mean average error of about eight meters, which subsequently leads to irregular speed calculations. Furthermore, the GPS signal can be even less accurate on cloudy days, in the vicinity of skyscrapers, or in tunnels. To minimize the influence of erroneous data values, methods like Kalman filtering can be utilized. A Kalman filter is based on a time-series of data values and estimates the current value using a probability distribution and dynamically adapting parameters [13]. Regardless of the exact information source and processing algorithms, the vehicle speed is an important metric in almost all driving data acquisition systems.

OBD-II Interface Adapters

Previous work shows that smartphones are not the only possible source of driving data in low-cost setups. On-board diagnostics (OBD) adapters are a complementary way to gather vehicle data. The term OBD refers to a vehicle's self-diagnostic and reporting capabilities. OBD-II is an improvement over the initial standard and specifies a connector interface and a set of command codes. In general, car models released 1996 or later are compliant with the OBD-II standard. Adapters for this interface provide direct access to the internal vehicle system status, yielding readings such as accurate speed, fuel consumption, pedal pressure, or steering wheel angle. OBD-II dongles were originally developed for maintenance and diagnostic purposes, however, since they often implement a Bluetooth, WiFi or cable-based interface, they can be leveraged for real-time driving data collection and analysis. This connection is especially useful as it allows acquiring accurate speed information independent of the often inaccurate GPS signal. OBD-II devices have been used to complement data by projects investigating low-cost setups (e.g. in [7, 44, 45, 8]) as well as in traditional NDS such as [4].

Commercial consumer systems that use OBD-II adapters come at an affordable price (below 100$) and track the driver's activity and skills. Using the smartphone apps that are part of the product such as that of *Automatic*[4] or *WayRay*[5], drivers can visualize their trips and get coaching on how to improve their driving behaviour, e.g. how to save fuel (see Figure 3.4). The systems recognize start and end of trips automatically and, once the adapter is installed in the device, do not require any user interaction to function correctly.

(a) Automatic (b) WayRay Element

Figure 3.4: (a) Automatic's OBD-II dongle and smartphone app, (b) WayRay Element *Autoyoga* driving coach

Several studies compared the *accuracy of data* collected by smartphones to data acquired through OBD-II connectors. Meng et al. [46] suggest fusing smartphone sensor information to reach a similarity of 96% between the two data streams. Sathyanarayana et al. [45] report a higher accuracy in recognizing driving maneuvers when using only smartphone sensors compared to only OBD-II data. Even though these results suggest that an OBD-II adapter is not necessary to increase the accuracy of certain data values, other values such as vehicle speed are known to be erroneous

[4] www.automatic.com, accessed 03.08.2016
[5] www.wayray.com/element, accessed 03.08.2016

when sensed by a smartphone [44]. Moreover, specific information such as fuel consumption or pedal use cannot be acquired using a smartphone alone. For that reason, OBD-II devices remain a widely used addition to data acquisition systems.

Vehicle Data Processing

Based on raw data values such as acceleration, rotation and speed readings, there are different approaches to process vehicle data in order to detect critical driving events. The simplest way to process the data stream is to define a set of **thresholds** based on previous empirical observations. Once the sensed values exceed a threshold, an event notification is triggered. The app *DriveSafe* relies on a set of thresholds to differentiate between low, medium and high amplitudes in acceleration and orientation data [42]. Another example of the use of thresholds is the project of Dai et al. [47] where the mobile phone is programmed to detect drunk driving based on speed control and lane positioning. It should be noted that thresholds for events are usually based on empirical observations. Consequently, variations can be found in different implementations (e.g. between [42] and [48]). Nonetheless, thresholds are a good compromise between complexity and accuracy in detecting maneuvers and events.

A more complex technique is **dynamic time warping (DTW)**, where the data of a specified time frame is compared to a set of templates. If the data is sufficiently similar to one of the templates, the system concludes that a certain driving maneuver occurred. The key benefit of DTW is that the algorithm can detect maneuvers independent of their amplitudes or durations which allows matching data from different drivers to the same template set. Johnson et al. [5] implemented a DTW algorithm that classifies twelve different types of events including turns, accelerations and excessive speed. To detect the start- and endpoint of an event, they compared a predefined threshold to the simple moving average of raw data values. The research team reports an accuracy of 97% when detecting aggressive events. A similar technique was applied by Eren et al. [49]: Using DTW to detect maneuvers they performed a bayesian classification to rate the drive as either *risky* or *safe*. While DTW is a more sophisticated approach to maneuver detection than simple thresholds, it is still comparatively easy to implement and frequently used.

In more recent years, multiple research projects implemented **machine learning** algorithms to detect critical driving events in vehicle data. For this technique, the system has to be trained prior to the actual deployment to be able to distinguish between the different events or driving styles. Hong et. al [7] built a machine learning model to classify driver aggression based on vehicle movement data. Van Ly et al. [50] applied supervised (support vector machine) and unsupervised (k-means) learning algorithms on vehicle data to differentiate between different drivers. They found that braking and turning maneuvers show the highest potential in differentiating between drivers.

These three techniques come with notable benefits and drawbacks: Defining threshold values is easy to implement and very fast to compute but lacks the ability to adapt to driving styles. Furthermore, it is critical to choose the right threshold values for detection. Otherwise, a high rate of false positives or false negatives can be expected. DTW can reliably detect driving maneuvers for a single driver. However, providing a template set that detects maneuvers for different driving styles can be challenging. In addition, the processing time is higher than a detection based on simple thresholds. Even more taxing on the device's hardware are machine learning concepts, however, these techniques are the most flexible ones: a classifier can be trained to differentiate between different driving styles [50]. Consequently, it is possible to detect anomalies in a person's driving behaviour which potentially allows drawing conclusions about the driver's current mood (e.g. aggressiveness [7]).

In summary, vehicle data in the context of driving style constitutes any data values that can be used to describe the vehicle's movement and operation. Not all information on the vehicle can be measured by a single smartphone, which is why some previous projects chose to extend their system by using an OBD connection. Several approaches for maneuver detection exist, varying in complexity and flexibility.

3.2.2 Environment Data

The second type of data, environment data, describes not the vehicle itself but the environment around it. Two major classes constitute the driving environment of a vehicle: firstly, environment data includes information about other nearby road users, including vehicles, cyclists, and pedestrians. The traffic infrastructure such as the type of road, road markings, and traffic signs form the second class.

The phone's back camera is one of the main sources for information about the environment around the vehicle. When the camera is pointed at the road ahead, it is possible to apply image processing algorithms to extract information on following distance [51], lane deviation [52], and upcoming road signs [43].

Following distance is an important metric to understand the driving context and can be extracted through different image processing methods varying in complexity and accuracy. Kumtepe et al. [52] used a technique called *Inverse Perspective Mapping* (IPM) to detect road lines and the distance to the vehicle in front. For this processing step, four distinct points in the original image are chosen and mapped to four other points in a bird's-eye perspective of the same image. In the resulting top-down view of the road, the distances to other vehicles can be measured without distortions. IPM has also successfully been used in previous research projects to detect obstacles [53] and pedestrians [54]. A different method is to estimate the distance to the front vehicle by its pixel height (or width) in the captured image frame. This algorithm relies on accurate information about the vehicle's real-world width, the camera's sensor dimensions, and the focal length. The distance to the vehicle in front can then be calculated using:

$$distanceToObject(mm) = \frac{focalLength(mm) \times realWidth(mm) \times imageWidth(mm)}{objectWidth(mm) \times sensorWidth(mm)} \quad (1)$$

However, since not all vehicles have the same dimensions, this method is not always accurate, which is the reason why most smartphone-based research projects implement the IPM approach.

The **detection of road lines** is another important part of gathering environment data. In recent years, many different methods for lane marking detection were developed and evaluated in studies, mainly motivated by the rising interest in advanced driving assistance systems. The most popular approach is to apply a canny edge detector followed by the Hough transform algorithm for the actual detection (e.g. in [55, 56]). Once the positions of road lines are detected, the number of lane departures or the mean lane deviation can be used as a driving performance metric. A number of applications also provide real-time feedback to the driver in the form of lane departure warnings (LDW). For example, You et al. [51] display an alert icon in their application *CarSafe* when a lane weaving event is detected. Together with the following distance, the lane positioning constitutes a core part of environment data.

Lastly, information about the traffic infrastructure around the vehicle can be sensed in different ways. Advances in image recognition have made it possible to detect **traffic signs** and the status of traffic lights [43]. Going further, information about the traffic infrastructure can be shared with other vehicles to create a more meaningful context. This concept was implemented in *Signal-Guru* [57], a smartphone application that detects the state of traffic lights through image processing and shares the information with other users in order to predict the schedule of traffic lights. A similar application was developed by Apple et al. [58], however, instead of using image recognition, they rely on information provided by official infrastructure operators. In the future, combining multiple sources of environment data could be used to estimate upcoming safety risks in real-time, as envisioned by Andreone et al. [59]. These examples show that image recognition and vehicle-to-vehicle communication of crowdsourced data can greatly increase the understanding of the driving environment.

3.2.3 Driver Data

The category of driver data consists of all readings and processing results that allow deducing the current driver state. This includes the detection of driver distraction, drowsiness and physiological measurements such as the heart rate.

For the detection of distraction and drowsiness, most previous projects rely on **image-based approaches** (e.g. [6, 60, 61]). A majority of modern phones is equipped with at least two cameras, one located at the back of the device and one on the front next to the display. This fact enables researchers to apply image recognition algorithms not only on images of the road but also of the driver. In the case of distraction and drowsiness, researchers focus on the front camera's video stream.

The first step to extract information on distraction and drowsiness is a facial recognition algorithm. Implementations vary in complexity and robustness. More sophisticated systems then apply further processing steps to measure eye blinking frequency, gaze movement, yawn frequency, and head position.

A strong indicator of driver's fatigue is the **percentage of eye closure** (PERCLOS) [62], counting the occurrences of eye closures during a fixed time interval. If the PERCLOS score exceeds a certain threshold, it can be assumed that the driver is drowsy. Jo et al. [60] use this metric in their project and propose to combine it with a similar metric they introduced, PERLOOK, which expresses the percentage of eyes-off-the-road-time during a time interval. With this approach, their system can measure drowsiness and distraction simultaneously.

One major drawback of image-based analysis of the driver's state is that RGB cameras and the algorithms need sufficient illumination to work reliably. Professional hardware setups for driving studies try to avoid this problem by installing additional infrared cameras, which is not feasible for low-cost, single-device setups. Other solutions suggest adding external illumination by night [61] which may cause safety hazards due to blinding the driver.

A different source of raw data fit to extract the driver's current state are **physiological measurements**. In the driving context, several physiological measures have been used in the past. Electrodermal activity (EDA) refers to the variation of the electrical characteristics of the skin in response to sweat. This measure can be leveraged as an indicator of autonomic sympathetic arousal as EDA increases in situations of stress or high cognitive load and decreases in low arousal states [63]. The most basic physiological measurement is the heart rate; it can be sensed by electrodes on the skin (electrocardiogram, ECG) or by optical sensors that read changes in blood flow. Examples for the use of the heart rate in the context of driver data include drowsiness detection [64] and cognitive workload measurements [29]. Both EDA and heart rate activity can provide valuable information on the driver's arousal levels but are often used only in laboratory settings due to the size and complexity of the equipment.

The recent rise in popularity of fitness wearables such as the *FitBit* or smartwatches makes it now possible to measure the heart rate activity and even skin conductance in a non-invasive way. More advanced alternatives such as chest straps require additional effort to put on and are not suitable for everyday driving. While studies show that consumer-grade heart rate sensors on the wrist do not reach the accuracy of more sophisticated systems, researchers argue that their data output could be reliable enough to detect trends and patterns during a drive [35]. For example, Ros-Aguilar et al. [65] demonstrated a concept for leveraging a smartwatch to detect drowsiness.

However, wrist-worn wearables are not the only non-intrusive way to acquire physiological data: A different kind of wearable is leveraged by Warwick et al. [66] in their study on drowsiness detection. They use *BioHarness 3*, a sensor by Zephyr that can be placed close to the body by wearing either a chest strap, a holder, or a special shirt (see Figure 3.5).

While this solution is currently targeted at professional customers and research, it demonstrates that more accurate measurements can be acquired without the need for a complex setup. Another example of how to measure the heart rate unobtrusively is presented in [59] where researchers equipped the steering wheel with an electrode fabric. If both hands are placed on this fabric, the system is able to sense the driver's heartbeat and send it to a connected smartphone.

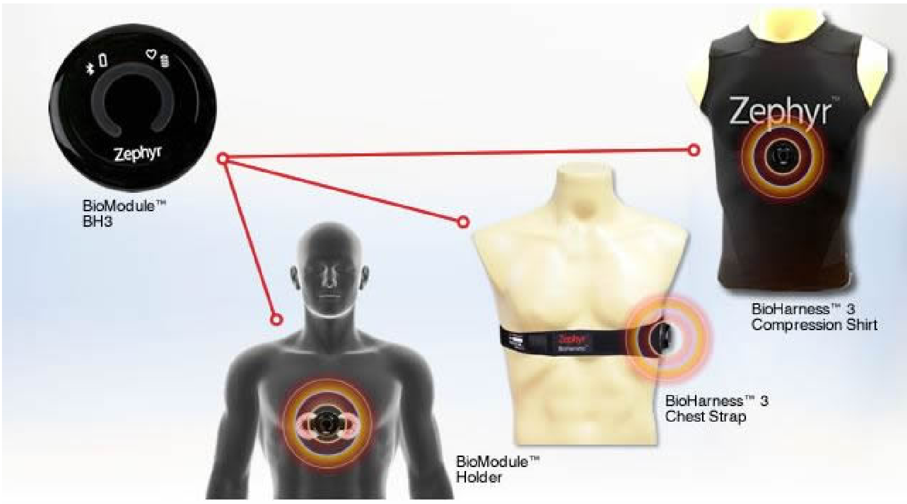


Figure 3.5: Bioharness 3 by Zephyr [6]

All in all, sensing driver data can be achieved through image recognition and physiological measures. Hardware in smartphones has become powerful enough to analyse image data in real-time to detect distraction and drowsiness. The two most common physiological measurements are heart rate activity and skin conductance. One key factor of physiological sensors is how invasive the data collection process is. While professional equipment is often heavyweight and expensive, recent consumer-grade fitness wearables include advanced physiological sensors of small size.

3.2.4 Combined Data

Often, research projects aim to combine vehicle, environment and driver data to create a complete understanding of the situation. In most cases, the cause of a change in one type of data can be found in another category: If, for example, the driver is distracted by his phone (=driver data), it is likely that the vehicle movement patterns will change (going above the speed limit, lane weaving, etc.).

Speeding detection is commonly implemented in event detection systems of mobile platforms. Such algorithms must combine the vehicle's current speed and location (=vehicle data) with information about the traffic infrastructure (=environment data). In order to acquire the current speed limit, both image-based road sign detection [43] and queries to a web database [67] have been used in previous projects. It should be noted that each approach has considerable limitations. Image-based detection is dependent on good illumination and fails if the traffic sign is occluded. A query to a database is only successful if the GPS signal is accurate and the database actually contains a valid value for the specific road. Due to temporary changes (e.g. construction sites), data entries

[6] https://www.zephyranywhere.com/products/bioharness-3, Accessed 08.09.2016

for speed limits are not always up to date. While hybrid systems combining both approaches have been developed and are deployed in higher-end cars[7], to this date, no smartphone-based hybrid solution has been described in previous work.

Another example of combined vehicle and environment data is the **time-to-collision** for which the current speed and the distance to the front vehicle must be known. This metric has been implemented in previous projects such as *iOnRoad*[8] (see Figure 3.6).

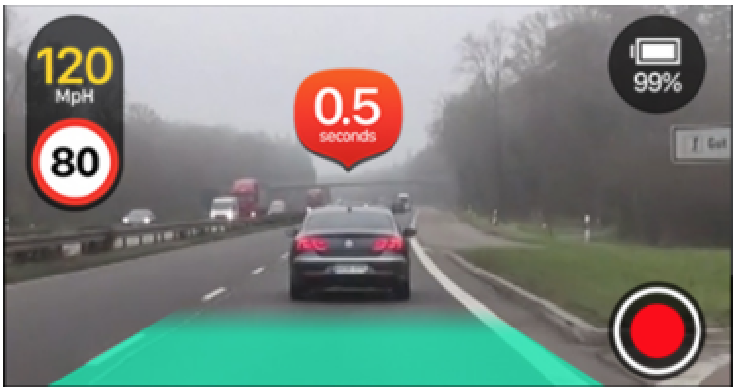


Figure 3.6: The smartphone application *iOnRoad* calculates the time-to-collision.

Recently, **driver aggressiveness** has become a popular area of research. While in earlier studies researchers had to rely on data gathered in simulators or through self-reports, newer smartphone-based systems allow judging the aggressiveness of driving behaviour on the real road. Hong et al. [7] report on a setup consisting of a smartphone, an OBD-II adapter, and an additional IMU. They extract the necessary information out of vehicle data to classify a driver as aggressive or calm. In contrast, Kumtepe et al. [52] fuse vehicle and environment data to detect aggressive driving. In their approach, the system monitors sudden lane changes, tailgating behaviour as well as speed and accelerations. Based on these events, a machine learning algorithm rates the level of aggressiveness driving style. However, while Hong et al. tested their system in a real-world study (N=22), Kumtepe et al. only did an analysis based on preexisting data from the 100-car NDS. A more thorough evaluation is needed to prove the benefit of combining environment and vehicle data for aggressiveness detection. In theory, adding more context to the data will improve the machine learning classifier's accuracy, as stated in [7, p.4055]: "We suspect that variables such as the weather and traffic or even the body position of the driver, could improve the classification accuracy by allowing the model to understand some intentionality behind behaviors". This notion shows that there is still untapped potential in sensor-based detection of driver aggressiveness, and thus in context-aware systems in general.

3.3 Naturalistic Driving Studies

As described in Section 2.2.3, Naturalistic Driving Studies (NDS) are a research method to collect data on real-world driving over a longer period of time. This method allows to better understand daily driving and to evaluate how drivers behave under natural conditions. Researchers assume that the presence of the data acquisition systems (DAS) has little influence on drivers after they get used to it [31]. With the advances in sensing technology, NDS have become more and more popular in recent years.

[7] http://www.bmw.com/com/en/insights/technology/technology_guide/articles/speed_limit_info.html, Accessed 14.09.2016

[8] https://play.google.com/store/apps/details?id=com.picitup.iOnRoad.pro, Accessed 14.09.2016

NDS are often conducted on a national or international level: Australia, The United States, Sweden, Europe all have already concluded or ongoing real-world driving study programs. The research focus of these large-scale studies is often broad, whereas smaller NDS focus on more specific research questions such as the effect of passengers on driving behaviour [68]. In the following, several examples for both large-scale and smaller-scale NDS are presented in more detail.

The so-called *100-car Naturalistic Driving Study* [4] was the first large scale NDS in the US starting in 2003, operated by the Virginia Tech Transportation Institute (VTTI) and the National Highway Traffic Safety Administration (NHTSA). Over the period of one year, driving data of 100 drivers, 2.000.000 miles and 43.000 hours was collected. The primary goal of this study was to identify crash causes and near-crashes. To understand what factors lead to a crash, the vehicles were equipped with a complex hardware setup, consisting of various sensors, radars, and five video streams. Table 3.1 shows a detailed description of the captured data. The *100-car NDS* also featured an "incident push button" which participants could press to flag a particular incident in the data collection.

One of the main results of the data analysis is that driver inattention is a more important issue than previously believed. Distraction often remains unmentioned in police reports but is one of the major factors leading to critical incidents: the collected data showed that in 78% of all recorded crashes the driver was not paying attention to the traffic [69].

The biggest European naturalistic driving study is *UDRIVE* [70, 71], which started in 2012 and is scheduled to end in 2017. Its research aim is to study driving behaviour in different European regions, especially in regard to the effects of driver distraction and the interaction between drivers and vulnerable road users such as pedestrians and cyclists. Another research aim of the study is to gain more insight into everyday driving. In contrast to the *100-car NDS*, *UDRIVE* does not only collect data from passenger cars but also from trucks and motorbikes. In total, 200 vehicles were equipped with the specifically developed DAS which has a strong emphasis on video data: depending on the vehicle type, up to eight cameras are installed. Figure 3.7 shows the angle and direction of the video cameras used.

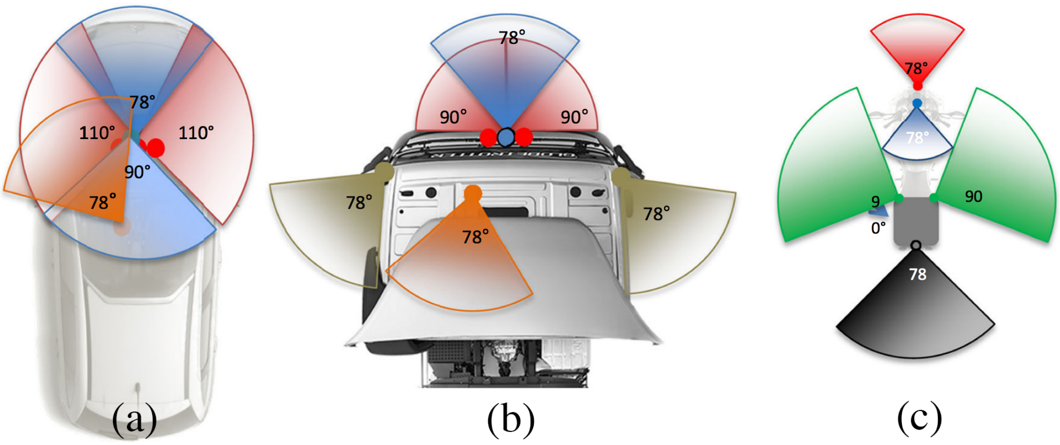


Figure 3.7: The video camera setup of *UDRIVE*'s data acquisition system for (a) passenger cars, (b) trucks and (c) motor bikes.

Similar examples of large-scale NDS are the *Australian Naturalistic Driving Study (ANDS)* and the *Sweden-Michigan Naturalistic Field Operational Test (SeMiFOT)* [48]. The *SeMiFOT* project differs from the other approaches as it was not only designed to monitor everyday driving behaviour but also to provide a platform for the evaluation of new safety systems.

Another research question of interest is the behaviour of **novice and teen drivers**. Simons-Morton et al. [68] conducted a smaller-scale NDS (43 vehicles, 18 months of data collection) to study the effects of passengers on risk taking and crashes/near-crashes among teenagers. Their findings show that novice drivers take more risks if peer passengers are present compared to adult passengers. A similar research focus is reported by Carney et al. [72]. In this project, the data collection was not performed by the research group itself. Instead, the authors analysed almost 1.700 teenage driver's crash data sets recorded by Lytx's DriveCam system[9]. The database consists of video, audio, and accelerometer data around critical driving events.

Table 3.1 provides a selection of existing large scale NDS, detailing the research aims as well as the extracted data. Some of the gathered data values are present in all of the listed NDS (e.g. accelerations (=g-forces)) and can therefore be seen as standard data while other areas of information are relevant to specific studies only (e.g. alcohol sensor). In summary, very large scale NDS explore a more general research focus and try to collect as much data as possible whereas smaller scale studies only collect a handful of values necessary to investigate a specific research question.

Study	**Period of Data Collection**	**Research Focus**	**Extracted Data**
100-car NDS [4]	2003–2004	Characterization of chrashes and near-crashes; Driver inattention	Longitudinal and lateral acceleration and turning rate; Vehicle Network Box: speed, brake, throttle, turn signal; Headway detection system to detect leading or following vehicles (radar); Side obstacle detection (radar); Incident box where drivers can flag incidents; Video based lane detection; Five cameras (driver, forward, rear, passenger, driver's hands) + IR illumination; GPS; Microphone; Automatic Collision Notification
UDRIVE [71]	2012–2017	Crash causation and risk; Driver inattention; vulnerable road users; driving style and eco driving	Acceleration and turning; video environment recording; GPS; sound; video data from 7 cameras; CAN data
Simons-Morton et al. [68]	2006–2008	effect of passengers on risky driving in young drivers; crashes and near-crashes	G-forces (hard braking, rapid starts, hard turns); GPS; video data;

[9]www.lytx.com, Accessed 15.09.2016

ANDS [10]	2016–2017	collision avoidance; distraction; inattention; speeding; aggression; tiredness	G-forces (hard braking, rapid starts, hard turns); number of passengers; lane position (mobileye); distance to vehicle (radar); GPS; video data (roadway, driver gaze direction, head tracking); alcohol sensor; temperature inside cabin; CAN data; sound level;
Fitch et al. [73]	2011	cell phone use while driving	G-forces (hard braking, rapid starts, hard turns); lane position; distance to front vehicle (radar); GPS; video data (roadway, driver, vehicle console); alcohol sensor; ambient light; OBD2 data (speed, braking, steering wheel); sound level;
SeMiFOT [48]	2008–2009	crash causation; effect of new safety systems	Event triggers for g-forces (hard braking, rapid starts, hard turns); MobilEye; GPS; video data (roadway, full internal view, driver face) w. IR illumination; CAN data (speed, braking, steering wheel); driver inattention (eyetracking)
AAA Teen NDS [72]	2007–2013	Young drivers distracted driving (focus on peer passengers and mobile phones)	Event triggers for g-forces (hard braking, rapid starts, hard turns); 12 seconds of video and driving data, 8 before, 4 after event; video data (front headway, driver);

Table 3.1: Overview of naturalistic driving studies, their research focus, and the collected types of data.

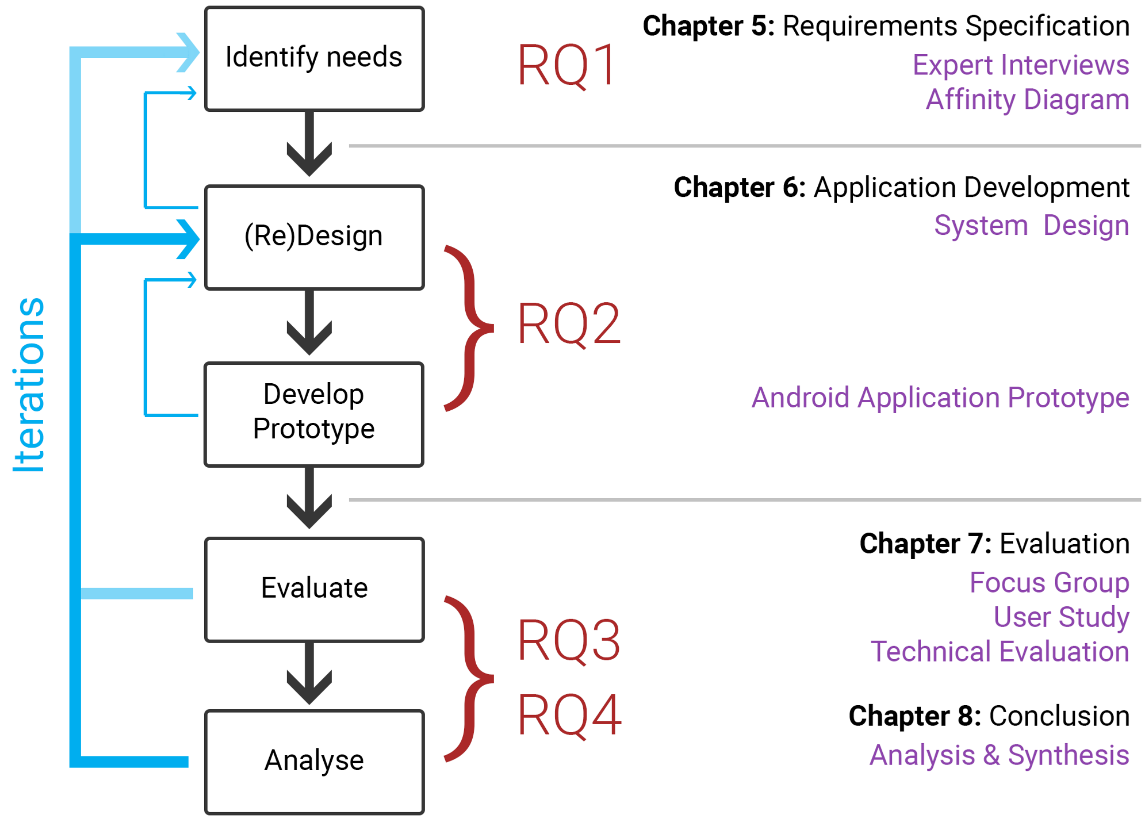
Summary
In conclusion, this chapter presented previous work in the areas of **activity tracking**, smartphone-based **driving style analysis**, and **naturalistic driving research**. As a prominent example of activity tracking, the quantified self movement is driven by mobile technology. Smartphone sensors and wearables allow individuals to track various aspects of everyday life and identify areas

[10] http://www.monash.edu/muarc/research/research-areas/transport-safety/australian-naturalistic-driving-study

of improvement. Most of the data collection process is done automatically by context-aware algorithms that detect and quantify activities. These findings directly relate to the area of driving style analysis, a field of research that in recent years also leveraged mobile technology. By building on the ubiquity of smartphones, driving data can be collected in a lightweight and large-scale way that was previously impossible. The three categories of data needed for driving style analysis, namely vehicle, environment, and driver data are also the main sources of information for naturalistic driving studies. Results deduced from naturalistic data promise a high real-world applicability and validity and are therefore an important complement to more traditional road safety methods such as surveys or simulator studies. The review of related work showed that most existing smartphone applications target the individual driver, not road safety researchers. This lack of a lightweight tool for experts presents the research gap targeted in the thesis.

4 Methodology

This project's methodology loosely follows the **interaction design model** proposed by Preece et al. [10]. Their model describes the life cycle of interaction design projects in a simplified form; it is an iterative process which encourages a focus on the user. Consequently, Preece et al.'s interaction design model has its roots in user centered design as described by Norman and Draper [74]. In Preece et al.'s model, a project starts with identifying needs and requirements. Based on this first activity, several alternative designs are generated to meet the requirements. These early designs are then implemented as interactive prototypes, which are subsequently evaluated in a user study. Building on the feedback gathered during evaluation, the next iteration of the lifecycle begins. Through the iterations, the design is consistently refined until it meets the initial requirements.



Research Questions

RQ1: What requirements do road safety experts have for naturalistic data acquisition systems?

RQ2: How can we design a smartphone-based data acquisition system in a way that meets the requirements of road safety experts?

RQ3: **a)** How do road safety experts judge the low-cost system in terms of usability?
b) How does the smartphone-based system perform in terms of accuracy?

RQ4: What are the benefits and disadvantages of a low-cost data acquisition system?

Figure 4.1: Methodology of this project, following the model from Preece et al. [10]. On the left side, the five stages of one iteration are depicted; the right side specifies the corresponding parts of this thesis.

In this thesis, one full iteration was completed, including smaller intermediate iterations between individual phases of the project, e.g. going back from the application development to redesign a specific part. The five main steps of the process as visualised in Figure 4.1 were as followed:

Requirements Specification: Following a comprehensive literature review, eight semi-structured interviews with road safety experts were conducted in order to identify the requirements of the proposed research platform. A content analysis [75] and affinity diagramming [76] guided the process of analysing the answers. These activities led to the definition of an abstract model and ultimately to the specification of a comprehensive list of functional and non-functional requirements. This list represents the answer to the first research question (RQ1).

System Design: On the basis of the previously defined requirements, a system architecture was developed and represented as a UML model. Furthermore, a number of paper sketches and storyboards were created to explore different options for the user interface and interaction with the system. Both the system's architecture and the user interactions were part of the activities around RQ2, i.e. the question about how to best design a system that fits the previously identified requirements.

Implementation: During this phase, a fully functional prototype was implemented in the form of a smartphone application. As part of RQ2, the implementation follows the previously defined system design and matches the requirements set by literature review and expert interviews. However, due to the prototypal nature of the implementation and the general scope of this project, some features specified in the list of requirement were excluded from this iteration.

Evaluation: The fourth project phase focused on answering the third research question (RQ3). Hence, this phase looked at two different aspects: Firstly, the usability of the system was evaluated in an expert user study which focused on the initial study setup and the connection of external devices. The second part in form of a technical evaluation was conducted to assess the accuracy and reliability of data collection. In order to achieve this, the developed research platform was compared to a professional DAS on multiple drives on the real road.

Analysis: The last step was to analyse the previous steps – from the requirements to the evaluation data – in order to extract the information needed to answer RQ4. Thus, this final phase synthesises all the findings to draw a conclusion and identify the areas which can be improved upon in the following iterations.

In summary, the presented project follows the interaction design model by Preece et al. [10]. The iterative process described in this model is formed of four major phases, with intermediate iterations between those steps. In an adaptation of the model, the process of this project included an additional fifth step. This thesis reports on one full iteration consisting of requirement specification, system design, implementation, evaluation, and analysis. Each of these stages helped to answer one of the four central research questions of the thesis. Putting all together, the findings of the last phase form the basis for a potential next iteration and may inform future research.

5 Requirements Specification

The thesis' main goal is to explore and build a research tool that enables road safety researchers to conduct inexpensive real-world driving studies. Several high-level project aims were included in the task description of this project:

- Modular design for simple extensibility
- High usability: no prior training needed, simple setup
- Focus on creating a "toolbox" instead of developing new algorithms

The final set of requirements that match these aims was derived from two different sources: A review of existing projects in literature (discussed in Chapter 3) and interviews with road safety experts. The next subsections present the methods and results of the requirements analysis phase, concluding with a comprehensive list of all requirements.

5.1 Expert Interviews

Road safety experts are the targeted user group for the proposed research platform. As described in Chapter 4, this thesis follows a user centered design approach. Consequently, interviews with road safety experts were conducted prior to the system's specification to establish the user's needs. The main research objective was to get an understanding about what constitutes naturalistic driving research as well as what features and properties a low-cost system should include.

5.1.1 Participants

A total of N=8 experts (6m, 2f, M_{Age} = 39.25, SD_{Age} = 8.58) were participating in semi-structured interviews. For recruitment, opportunity sampling [77] in combination with the key informant technique [78] was used. All experts are professional road safety researchers with prior experience in qualitative and quantitative data collection methods in the context of vehicles. Their educational background and research range from psychology to computer science (see Table 5.1).

ID	Age	Gender	Primary Area of Research
P1	39	male	Driver sleepiness and its effects on driving performance
P2	58	male	Interventions for older drivers
P3	36	male	Impact of driving apps on driver behaviour
P4	40	female	Secondary data sources (police, hospitals)
P5	32	male	Safety at level crossings
P6	31	male	Driving simulators and naturalistic driving studies
P7	35	male	Microscopic traffic flow model, interactions between road users
P8	43	female	Behaviour and performance of teenage drivers

Table 5.1: Details about the interview participants.

Before commencing data collection, the procedure was approved by the university's ethics committee (approval number 1600000713, in accordance with the *Australian Code for the Responsible Conduct of Research*) and written consent was obtained from participants. Each interview lasted approximately 30 minutes. Audio recordings ensured the focus on the participant's answers instead of taking detailed notes.

5.1.2 Interview Objectives

The purpose of the interviews was to inform the development of the research platform by expert users. Thus, the following three research objectives were defined prior the the interviews:

(a) What kind of driving data are experts interested in?

(b) What are important critical driving events?

(c) Which are the core aspects of driver behaviour?

The questions (see Appendix A) were designed to start with a discussion about the expert's current research affiliation, followed by detailed questions about data collection and more general questions to explore the broader context of this project. The experts were encouraged to create a "wishlist" of data collection features they would like to use in real-world driving studies. The interview concluded with questions about future challenges in road safety research.

5.1.3 Analysis

The analysis of interview data was done in three steps. First, a simple form of content analysis [75] was performed. All topics mentioned by the road safety experts were written on sticky notes and grouped into three basic categories:

1. Research interests
2. Data values (driver and driving behaviour, vehicle data)
3. Concerns

Figure 5.1: Affinity diagram created to structure participant responses. Color coding: *(Light Green)* Research interests, *(Blue)* Data values, *(Magenta)* Concerns, *(Yellow)* Sensors

These three categories were then structured in the second step using an affinity diagram (see Figure 5.1). Affinity diagramming, also referred to as the "KJ Method" [76], was first described by Jiro Kawakita and is a technique to cluster a large amount of notes by grouping those with similar intents, problems, or issues. It helps to understand the relationships and connections between single answers. During the process of creating the diagram, four clusters of interview answers emerged:

- Everyday driving
- Information on the driver state
- Context around the vehicle
- Vehicle movement

In the last step of the analysis, an abstract model was developed based on the previously extracted information. This model (see Figure 5.2) represents the relations between clusters and other themes identified in the interview data. In the following paragraphs, the information collected in the interviews is presented top-down, starting with the abstract concept of what constitutes everyday driving, then discussing events and the raw data sources mentioned by the experts.

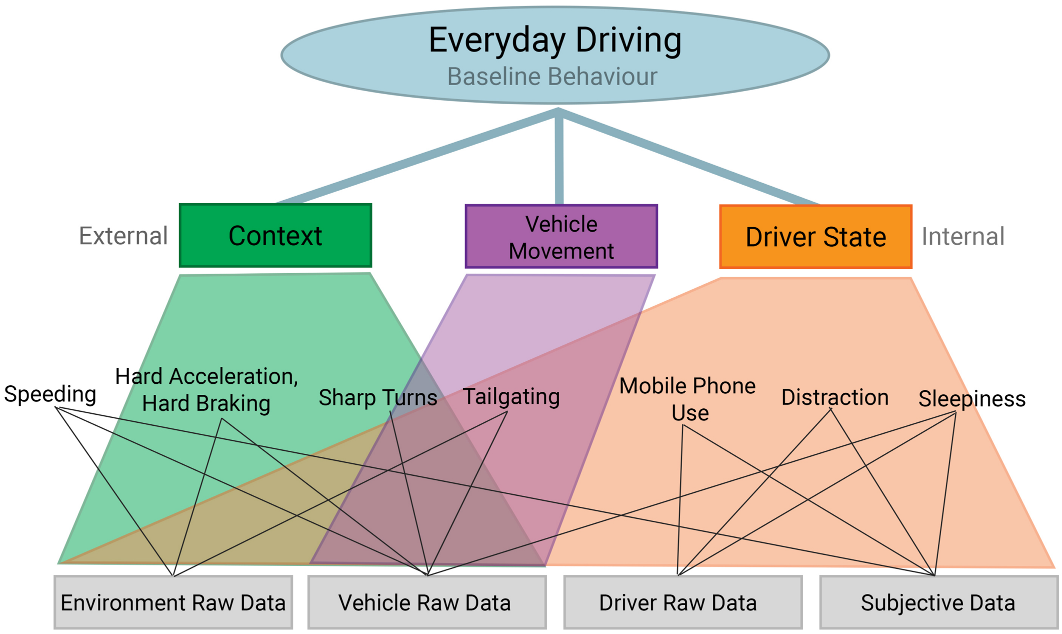


Figure 5.2: Abstract representation of the interview data: Information about a driver's everyday driving behaviour comprises the general *driving context*, a profile of *vehicle movement* on routine drives, and the *driver state*. These categories rely on different types of raw data, namely the ones presented in Section 3.2: environment data, vehicle data, and driver data. In addition to that, subjective data acquired through surveys can complement the information about everyday driving.

Everyday driving

A central goal of naturalistic driving research is to understand how drivers behave in normal situations. According to the experts, large-scale data sets collected over a longer period of time allow extracting the baseline driving style of a person. Building on this information, they can then identify situations in which something out of the ordinary happened (such as a crash) or extract time segments where the driver acted differently to his usual behaviour. A complete understanding of everyday driving requires information about the driving context, the movement pattern of the vehicle and the driver's state.

Driving Context
As part of everyday driving, most of the interviewed experts stressed that it is very important to get an understanding of the current driving context. This includes the position and behaviour of other vehicles, cyclists and pedestrians as well as traffic infrastructure such as traffic lights, road signs or level crossings. In addition to that, the experts were interested in factors that describe the details of a drive: the trip length, time of day and weather conditions. When asked about what context data they wanted in an ideal case, interview participants talked about the concept of vehicular communication systems. Connecting vehicles, pedestrians and infrastructure to exchange real-time information would allow road safety experts to study the causes and effects of driver and driving behaviour in more detail. Such a system "allows you to see [...] five hundred meters around your car and get the position of every other vehicle [...], then you could look at the behaviour of the driver inside that contex" (Participant 6). The underlying values (position, speed, intention of other road users) can be categorised as environment data, which forms the first source of raw data informing the driving context. The other source of raw data is vehicle raw data, i.e., the current position of the own vehicle, the speed, and other values to make more sense of the situation. In summary, environment and vehicle raw data build the basis for an understanding of the driving context.

Vehicle Movement
The second category that helps road safety experts understand everyday driving is the movement of the vehicle. All road safety experts participating in the interviews identified tracking the movement data as a key part of naturalistic data collection. This mostly consists of learning how a driver handles a vehicle, in particular how he accelerates, brakes and turns. In contrast to the driving context, vehicle movement information is based on only one category of raw data, namely vehicle raw data. The knowledge of vehicle movement is tightly linked to a good understanding of the current driving context, as described previously.

Driver State
Lastly, information on the behaviour and state of the driver was another prominent interests occurring in the interview answers and represents the third category of everyday driving. Among research interests in real-world driver behaviour, sleepiness, stress and distraction were named as major areas. Sleepiness and stress or arousal levels are dependant on physiological information, whereas distraction can be studied from different sources. Participant 2 highlighted the importance of identifying and understanding distraction: "The naturalistic driving studies show that distraction is a far more important issue than we thought it was." Answers mostly specified two different sources for driver distraction: in-vehicle distractions and events happening outside the vehicle. In-vehicle distractions are caused by the infotainment system, passengers or – most often mentioned – mobile phone use. Experts stated that any kind of distraction is of interest because of its major effects on driving behaviour.

Specific raw data values collected by sensors were also discussed during the interviews. Independent of the exact type of raw data, Participant 1 stressed how important it is to have a "clear documentation of sampling rates and things like that". A reliable data analysis after the study is only possible if the exact parameters such as the sampling rate are known. Finally, experts gave suggestions on how to gather specific raw data that is necessary to understand the context, movement and driver state.

Environment Raw Data
Current context-aware systems in vehicles rely on cameras and radars to detect and identify objects around the vehicle. Experts named the following distance or time-to-collision as an important metric of this category. Going further, any interaction with other road users are of interest to the experts. However, using cameras and radar as primary sources for information on the context does

not create a complete picture. For this, inter-vehicular communication would be necessary. Until this technology is ready to be deployed on a larger scale, information fusion from cameras, radars and other sources can provide valuable but less complete information to road safety experts.

Vehicle Raw Data

As stated before, vehicle raw data informs each of the three aspects of everyday driving: the driving context, vehicle movement patterns and the driver state. Experts suggested in particular to include the acceleration and the GPS signal. These two data sources provide measurement of speed, location and driving style. Based on the vehicle raw data, driving events such as hard acceleration or brakes can be automatically detected. The GPS signals allow experts to understand not only when but also where these events occur, which is important because some road segments, e.g. intersections or roundabouts, are more critical than others: "Merging is a [...] huge problem and highly stressful for young drivers" (Participant 8).

Driver Raw Data

For the following areas describing the driver state, specific metrics were mentioned by the experts.

- **Sleepiness**: There are several different ways to infer whether the driver is getting sleepy. The driver's blink duration was mentioned as an important measurement for detecting fatigue. This technique requires either a RGB video camera pointed at the driver's face or a more specialised approach with infrared equipment. In addition, the cabin light level derived from the smartphone's light sensor can provide additional information valuable for detecting sleepiness. To complement this data, the steering wheel reversal rates can be read from the OBD-II adapter[11]. As shown by Krajewski et al. [79], information about the steering wheel is an unobtrusive yet accurate way to infer driver drowsiness. Similarly, one expert stated that the use of pedals can indicate certain driver states, for example drunk or drugged driving.
- **Arousal**: Arousal levels range from monotony on one end to stress on the other end. As stated, one way to collect arousal data is the use of a physiological measurement devices such as a fitness wearable on the chest or wrist. For motorbikes, the ambient sound level could indicate the presence of a highly stressful environment.
- **Distraction**: Experts strongly wished a way to detect events such as an incoming phone call and how the driver reacts to this kind of interruption. In more general terms, several experts were interested in how a driver allocates his time and attention. Eye trackers could help to monitor and quantify this aspect of driver behaviour: "They spend more time taking their eyes off the road than they actually realize." (Participant 2). Sources of distraction outside the vehicle are harder to predict and can therefore not be as easily studied with specific sensors. Videos of the surroundings were named as the most promising source of information on this type of distraction.

Subjective Data

To complement the objective driver state data collected by sensors, two experts suggested that it would be valuable if the system included a "driving diary", asking the driver to complete a short questionnaire after each trip. The questionnaires could be used to compare the driver's own perception of the drive with the raw data values to identify and further examine disparities. Moreover, a driving diary could be used to gain more general information about a drive, e.g. the reason or motivation for the trip and whether the driver enjoyed it or not.

Concern: Accuracy

Several experts were concerned about the accuracy of the data collected by a smartphone, especially the GPS signal. Since the calculation of speed is based on GPS data, inaccurate position

[11]Note: It is not possible to read the steering wheel angle with all OBD-II adapters. This feature depends on the specific car model and OBD-II hardware used.

signals can lead to wrong speed values. Countermeasures such as Kalman Filtering exists but can not completely eliminate the error in the data. This is why participants suggested to optionally use an OBD-II adapter to access the vehicle's CAN bus. With that connection, accurate speed readings could be acquired, as well as information on fuel consumption, pedal use or the current state of headlights. The experts' concern about the smartphone's sensor accuracy can thus be addressed to a certain degree.

Similarly, one expert expressed concerns about defining a fixed set of thresholds. As described by Participant 8, "it's problematic if we're just going to choose thresholds" and "we actually need other complementary data. [...] You would need the visual to actually see if it is a problem". She suggested to fuse different data sources to compensate for possible issues with accuracy.

Concern: Data Storage
Another major concern of road safety experts regarding the presented research platform was the method of data storage and transmission. They identified several aspects as possible challenges:

- *How does the system know when to start data collection?*
  The system should only collect data when the car is driving. Therefore, the start and end of a trip must be reliably detected. Participants suggested to use either the accelerometer or the OBD-II connection to acquire the necessary information.
- *How do researchers get access to the data?*
  Researchers behind a naturalistic driving study need to be able to access the data in an effective way. Previous NDS usually established a wireless connection to a remote server to upload the data files. However, as two participants pointed out, a large amount of videos may require a WiFi connection for a reliable upload.
- *Where is the data being stored?*
  If a remote server is chosen as the central data storage, it must fulfill certain criteria in terms of security and accessibility. To protect participants' privacy, information has to be encrypted and transmitted securely. Additionally, researchers considered it as important to have unrestricted access to the raw data. Some existing off-the-shelf solutions for naturalistic driving research such as the VTTI system save the data on central servers and can restrict access to certain information.

These aspects highlight the researchers' concerns towards the handling and storage of data.

5.2 Final Requirements

Based on a combination of the project's general goal, the literature review and the expert opinions gathered in the interviews, the final requirements were defined. The following two subsections present the identified functional and non-functional requirements for the smartphone-based research platform.

5.2.1 Functional Requirements

Functional requirements (FR) specify all functions and desired behaviours of a system. In the case of the proposed sensor platform, this section includes the details regarding the sensors used for data collection and the critical driving events to be detected through processing algorithms.

FR 1 – Smartphone Sensor Data: *The system can collect and store data from the smartphone sensors as specified in Table 5.2.*

Section 2.1.1 provides an comprehensive list of all available sensors in modern phones. Out of these, only certain sensors are relevant for sensing real-time driving data. Table 5.2 lists the smartphone sensors to be used for data collection along with details on the exact data values, the unit of measurement and the desired frequency. The specified frequencies are either constraints set by technical limitations (GPS signal, image recognition) or informed by literature to fit the researchers' needs and algorithmic requirements (e.g. accelerometer frequency).

Sensor	**Data Value**	**Unit**	**Frequency**
Vehicle Raw Data			
Accelerometer	Longitudinal Acceleration	m/s^2	$\leq$ 20Hz
	Lateral Acceleration		
	Vertical Acceleration		
GPS	Location	Latitude, Longitude	1Hz
	Speed	km/h	
Rotation Vector	Yaw	rad/s, degree	$\leq$ 20Hz
	Pitch		
	Roll		
Driver Raw Data			
Front Camera	Face Detection	Boolean	$\leq$ 5Hz
	Blink Duration	Seconds	—
Microphone	Ambient Sound Level	Decibel	$\leq$ 5Hz
Environment Raw Data			
Light Sensor	Cabin Light Level	Lumen	1Hz
Back Camera	Following Distance	Meter	$\leq$ 5Hz
	Lane Deviation		
System	Time of Day	Time	$\leq$ 20Hz
	Outdoor Temperature	Degree Celsius	Every 5 Minutes
	Weather Condition	Description	

Table 5.2: The smartphone sensors for data collection along with the data values, the unit and the desired frequency.

FR 2 – Trip Detection: *The system automatically detects the start and end of a trip and controls data collection accordingly.*

Detecting the start and end of a trip needs to be trigger-based. Several options stand out for the actual trigger implementation. For a robust trip detection, a hybrid approach combining several triggers is preferable. The detection of a trip start and end must happen without any user interaction. While the system is waiting for a trip to start, battery consumption should be kept at a minimum and greatly reduced in comparison to the data collection process. The need for a robust trip detection is frequently described in previous work and was mentioned by experts in the interviews.

FR 3 – Parameter Customisation: *The system allows the researcher to activate individual sensors and customise parameters such as sampling rates and thresholds.*

The interviews made obvious that not every researcher is interested in the same aspects of driving data. For that reason, the system should allow researchers to select the individual sensors required for the specific study. Even more, the thresholds used for event detection must be customisable as they might differ between user groups or research aims. Finally, the system should provide the option to control the frequency of raw data output to allow for customisation based on the particular driving study. Recommendations in literature suggest values between 20 Hz and 1 Hz (see for example [5, 48, 80]). Each output of raw data values should be combined with the exact timestamp specifying when the data was measured.

FR 4 – External Sensor Data: *The system allows connecting to the external devices specified in Table 5.3 for increased data collection capabilities.*
As identified in the interviews, road safety experts are interested in some data values that can not be accessed by a smartphone. As a result, the developed sensor platform running on the phone must be able to connect to external sensing devices to increase the data collection capabilities.

Device	Data Value	Frequency	Price
OBD-II Adapter	Speed	1Hz	15$–35$
	Fuel Consumption	1Hz	
	Engine RPM	1Hz	
	Pedal Use	Event-based	
	Steering Wheel Angle	1Hz	
Fitness Wearable (e.g. FitBit Charge HR[12])	Heart Rate	Hardware dependent	50$–200$
	Skin Conductance		
Second Smartphone	Notifications, interaction with phone	Event-based	300$–500$

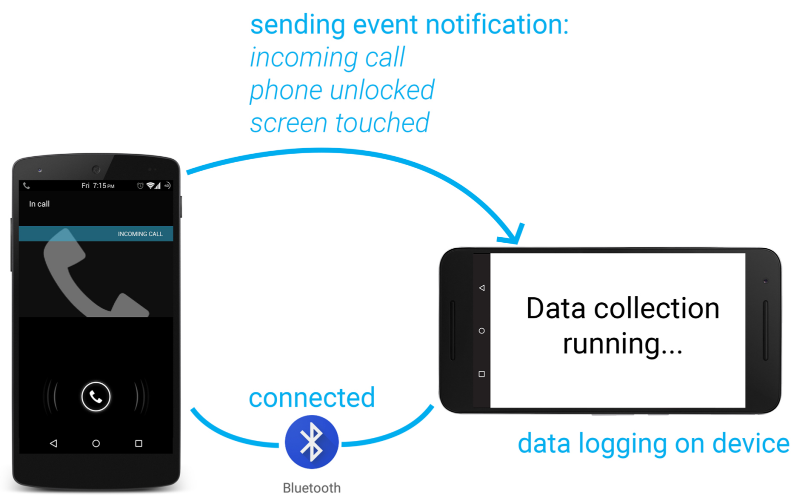
Table 5.3: Optional external devices to be supported by the application for increased data collection abilities.

FR 4a – OBD-II Adapter: *The system can connect to and process data from an OBD-II adapter.*
Several values of interest that are not accessible by smartphone sensors belong to the category of raw vehicle data. Among these, accurate speed, pedal use, fuel consumption or the steering wheel angle were named by experts more than once. This information can be acquired from an OBD-II adapter plugged into the car. Many adapters feature a Bluetooth interface, allowing to easily connect the smartphone to it. The usage of the external OBD-II device can be optional, depending on the particular study setup and research aims. This way, researchers have access to a broader range of data values but are not obligated to equip the vehicle with an additional device. Examples from literature show that the approach of an extendible system has been successfully implemented (e.g. [7]). The exact data values to be collected and sent by the OBD-II adapter are specified in Table 5.3. In addition to the broadened data collection capabilities, the inclusion of a direct connection to the vehicle also increases the accuracy of some values (e.g. speed). This addresses the experts' concern about the accuracy of smartphone sensors. This requirement was derived from both related work and expert interviews.

FR 4b – Fitness Wearable: *The system can connect to and process data from a wearable that has the ability to sense the heart rate.*
Arousal levels can be measured based on the driver's heart rate and skin conductance. The system should therefore be able to connect to a wearable physiological sensing device. As described in Section 3.2.3, different devices can be leveraged to sense heart rate data. However, the device to be used in the frame of this project has to satisfy certain constraints:

[12]https://www.fitbit.com/chargehr, accessed 26.08.2016

firstly, a key requirement for equipment in naturalistic driving studies is to be as unobtrusive as possible and secondly, the proposed sensor platform is designed to be a low-cost setup, ruling out professional physiological hardware. Accordingly, a consumer-grade fitness wristband or smartwatch should be used. Details of such a device are specified in Table 5.3. Physiological measurements have so far not been part of mobile sensing platforms but were wished for by the experts.



(a) User Phone (b) Research Phone

Figure 5.3: Setup with two smartphones: (a) the user's personal device, (b) the research phone provided by researchers.

FR 4c – Second Smartphone: *The system can connect to and process data from a second smartphone.*

The requirement to monitor phone interaction while driving was derived from the interviews. A recurring interest was if and how the driver interacts with his personal phone while driving, e.g. the reaction to an incoming phone call. However, with the application running on the user's smartphone, it is not possible to simultaneously use the cameras for image recognition and allow the user unrestricted usage of the device. This raises the requirement of extending the system: to reflect undisturbed real world situations, the study setup should consist of two separate smartphones. A dedicated "research smartphone" could be permanently installed in the vehicle, while the driver's personal phone automatically connects to it once the driver gets into the car. Through this connection, events such as incoming notifications and phone interactions can be logged on the research phone, complementing the collected driving data. Figure 5.3 shows a simplified diagram of the setup. The primary benefit of this method is that the driver can use his personal phone without any restrictions, resulting in natural data. A similar functionality is currently not part of commercially available DAS. Accordingly, implementing this requirement significantly increases the value of collected data to road safety researchers.

FR 5 – Event Detection: *The system can process and fuse raw data to detect and store driving events specified in Table 5.4.*

On the basis of the collected raw data, the system should implement event detection algorithms for the events specified in Table 5.4. This list of critical events was derived from both expert answers and previous work. For events that are based on thresholds, three levels of severity are specified.

Event	Sensors Used	Default Thresholds
Hard Acceleration	Accelerometer	0.2g, 0.3g, 0.4g [42, 48]
Hard Braking		
Sharp Turn	Rotation Vector (Accelerometer + Gyroscope)	0.45 rad/s, 0.6 rad/s, 0.7 rad/s [81]
Speeding	GPS, OBD-II, Map Service	Speed > Speed Limit
Tailgating	Back Camera, Speed	Time to collision. Below 60km/h: < 1s Above 60km/h: < 2s[13]
Distraction	Front Camera	Eyes-off-the-road > 3s
Mobile Phone Usage	Connection to second phone	—

Table 5.4: Critical events that should be automatically detected by the system.

FR 6 – Event Video Data: *The system can store videos of specified length showing the scene leading to and following a detected event.*

Videos allow analysing driving events from various perspectives and are as such a fundamental part of naturalistic driving research. Therefore, the research platform must implement the ability to save videos of critical driving events. This includes the seconds before and after the event occurred. The method of saving videos only around events instead of saving the whole drive reduces the amount of storage needed and shortens the time researchers have to invest in post-analysis. This approach has been successfully applied by previous studies such as the AAA Teen NDS [72].

FR 7 – Study Setup: *The system allows a researcher to create a new study and enter individual participant information.*

As a research tool for conducting real-world driving studies, the application must provide an interface to create a new study or select an existing one. In addition, researchers must be able to add participant information (such as age, gender, identification number) and link the participant to an existing study. Even though this requirement was not directly mentioned by experts or in literature, it is tightly coupled with FR 3 (parameter customisation) and was as such derived implicitly.

FR 8 – Questionnaires: *The system allows a researcher to create or provide surveys that are displayed to the driver after a trip.*

For the fourth category of raw data, subjective data, the system must allow road safety experts to define a set of survey questions. When the application detects the end of a trip, the questionnaire automatically appears on the phone's screen and asks the driver to complete the questions about the previous drive. This feature was mentioned in the expert interviews.

FR 9 – Data Storage: *The system stores the data on the phone and offers means to access the files remotely.*

As described in Section 5.1.3, data storage was one of the main concerns of interview participants, leading to the specification of this requirement. The first step is to save the raw data and video files on the phone. For this, the preferred file format are comma separated value (CSV) files. MPEG-4 or AVI should be used as the format for saved videos as they are widely supported by video players and other software for further processing. After the

[13] https://www.qld.gov.au/transport/safety/rules/road/distances/index.html, Accessed 18.08.2016

raw data values and video files are stored on the smartphone's internal flash memory, the system can handle the further procedure in three possible ways:

(a) Wait until a researcher manually downloads the files from the smartphone
(b) Send the files to a remote server over a wireless internet connection
(c) Let the participant review the data before uploading it to a remote server

Option (a) represents the simplest implementation of the working system and is the minimum requirement for this project. Option (b), an automatic file upload to a remote server, enables the researchers to monitor and analyse data even during the collection. As a result, problems and errors on individual devices can be identified early in the process which increases the time to find a solution. This option is preferred over the simple solution of option (a). Option (c) can be seen as an improvement over option (b), granting more rights and control to the study participants. While this is an useful feature if data privacy is an issue, it is not essential for the successful operation of a real-world driving study.

5.2.2 Non-Functional Requirements

Non-functional requirements (NFR) describe all aspects of a system that are not specific behaviours. Often, non-functional requirements are also referred to as "quality factors" and therefore specify criteria that can be used to judge the operation of a system. The following paragraphs cover data storage, data privacy, usability and extensibility.

NFR 1 – Data Privacy: *The systems encrypts any collected information on the phone and uses secure communication protocols to transmit the data.*

As pointed out by road safety experts, the collected data can contain sensitive information: besides the name, age and gender of a participant it shows where and when a person is driving regularly. Even more, the place of living of participants can possibly be extracted from the geospatial data. For this reason, it is important to specify and implement methods to prevent unauthorized access to the data.

This encompasses the data files stored on the phone and the transmission of files to a remote server. In order protect the files saved locally, there are two options: either to encrypt the data files itself or to generate a password protected folder containing the files. The transmission to the research server should be based on an encrypted communication between the smartphone and the server. In summary, the use of encryption on both files and communication ensures that no unauthorized person can read the sensitive data, protecting the participants' data privacy. As discussed under Option (c) of data storage, a complementary way to ensure data privacy is to allow participants to review their data before sending it to the researcher. Drivers could choose the trips and videos they are willing to share.

NFR 2 – Ease of Use: *The system can be operated by road safety experts without knowledge about the technology behind it.*

The targeted user group of the proposed application are road safety researchers. Many researchers in that field have a strong background in psychology and little to no prior experience with mobile platforms and the technology involved in collecting raw sensor data. Consequently, the setup of a study and the operation of it must be feasible without technical knowledge. This includes several aspects:

- The application should enable interactions through a simple and consistent user interface.
- Important settings such as thresholds and survey questions must be customisable without changing the application's source code.

- The collected data should be easily accessible to researchers through either a web-based application or a simple file manager on the phone.
- The connection to external sensors should be handled by the system without requiring a complex setup performed by the user.

NFR 3 – Extensibility: *The system's architecture allows for an easy addition of new components or exchange of existing algorithms.*

Even though the proposed sensor platform is designed to combine as many data values and sources as possible, new requirements to the software may arise in the future. The source code of the application should therefore be designed in a way that allows modifying the algorithms or add new external devices.

In conclusion, this chapter presented the process of defining the final requirements. To inform these requirements, eight expert interviews were conducted and analysed.

Thus, the chapter formulates an answer to the **first research question RQ1**: The analysis of interview data showed that "everyday driving" as baseline behaviour is a key interest of road safety experts in naturalistic driving research. Experts named a number of sensors and systems that can be leveraged in order to acquire this baseline data. In combination with a review of related work on mobile driving apps and naturalistic driving research, a set of functional and non-functional requirements was specified to match the interests found in the interviews. Apart from well-established data sources such as the accelerometer, new types of information on driver state (heart rate, phone interaction) were identified as requirement.

6 Application Development

This chapter informs the reader about the concept design and implementation of the described low-cost DAS. Building on the results of the requirements analysis and specification, a system design was developed and implemented as a smartphone app for the Android platform, further called *Sensor Platform*. The activities and decisions described here represent the efforts towards answering the second research question RQ2.

6.1 Limitations

Naturally, the approach of developing a low-cost alternative to existing data acquisition systems is facing several constraints. In order to be able to successfully employ the system in a real-world driving study, it is important to be aware of the existing limitations and their effects on data collection. This includes constraints that are inevitable due to the current state of mobile hardware as well as limitations set by the scope of this thesis.

6.1.1 Technical Limitations

The following list summarises the technical limitations that are set by the current smartphone hardware in comparison to instrumented cars of previous naturalistic driving studies:

- **Processing power**: Even though mobile phones have seen a substantial increase of processing power over the last years, they are not on par with modern desktop computers. For the developed application, mainly two factors are expected to be limiting: the central processing unit (CPU) and the main memory (RAM). Because the system collects and analyses many data sources simultaneously, some restrictions apply. For instance, related work indicates that the frame rate of processed images may have to be limited to a small number (2–5 FPS, e.g. [43]). Secondly, videos might have to be stored in a low resolution. Another expected tradeback is the reduced accuracy of some of the applied algorithms in comparison to an implementation on more powerful machines. These restrictions are necessary to maintain a real-time data collection process.
- **Number of cameras**: The majority of smartphones have two cameras installed: a front-facing and a back-facing camera. Even though these cameras are constructed with wide-angle lenses, they can not cover the same amount of perspectives as data acquisition systems previously deployed in NDS. For example, the *UDRIVE* NDS [70] equipped vehicles with up to eight cameras (compare Figure 3.7). The developed application uses only two cameras pointed at the road ahead and at the face and torso of the driver, leaving the rear- and side-views as well as passengers unmonitored.
- **Camera Positioning**: Related to the limited number of cameras, the exact positioning of the smartphone cameras is another constraint. The phone must be mounted in a way that keeps both the road and the driver inside the respective camera viewports. As a result, the alignment might not be as precise as would be possible with individually modifiable cameras which are traditionally part of a DAS.
- **Variations in video quality**: Besides the reduced resolution, the quality of videos from both front and back camera is affected by environmental factors such as light and weather conditions. RGB cameras installed in smartphones rely on a sufficient illumination to capture higher quality videos. As a result, videos of events happening at night will be too dark and might be affected by significant visible noise. The front camera is especially susceptible to bad lighting conditions since the vehicle's cabin is usually not illuminated while driving. Moreover, image recognitions algorithms such as the one to extract the following distance are expected to produce less accurate results under challenging weather conditions such as rain, snow, and direct sunlight, as explored by Rezaei et al. [82].

- **Data Accuracy**: Compared to high-cost data acquisition systems previously used in NDS, several data values are expected to be less accurate. One of these values is the estimated following distance which in professional DAS is often calculated based on a front radar. Since the proposed low-cost platform does not include radars, the headway can only be estimated by image processing techniques. Image processing is, as previously discussed, dependent on several aspects such as illumination and processing power. Another known source of inaccuracy is the GPS signal obtained by the phone. Errors in this type of data will not only affect the positional information but also lead to inconsistent speed calculations. In addition, the external physiological devices are known to have a high margin for error. Per requirement, the heart rate sensor must be unobtrusive, ruling out professional medical equipment. Consumer-grade fitness wearables deliver continuous but often inaccurate readings, as shown by Steinberger et al. [35].
- **Storage Capacity**: Modern smartphones models offer a storage capacity of up to 128GB. This can be limiting if a long data collection period is aimed for and no reliable data upload routine is defined. In general, the storage capacity is a limiting factor only if video recording is activated. However, the event-based video saving tries to minimize the issue as less storage is needed compared to the naive approach of saving continuous video data.

6.1.2 Delimitation of Scope and Excluded Features

In Chapter 5, a comprehensive list of requirements based on literature and expert opinions was defined. Due to the limited timeframe of this project, a subset of these requirements was implemented in the prototype of the first iteration. The resulting application is fully operational but does not include all specified features. However, the system's architecture is designed in a way that allows an easy extension of the platform, thus satisfying NFR 3.

Likewise, the focus of this thesis was not on developing new algorithms for detecting critical driving events. As a result, most implemented algorithms have known limitations, however, due to the open-source nature of this project, they can be easily replaced by more advanced techniques if needed. Among the excluded features, the following are the most notable and were left-out for reasons of complexity and low priority:

- The calculation of **lane deviation** is based on image recognition. Before the deviation can be extracted, the lane markings must be identified in each frame. This procedure adds another set of complex operations and would increase the required processing power. With further advances in computing power of mobile devices or a better optimization of the image recognition algorithms, this feature should be added in a future version of the platform. Related projects show the feasibility of an implementation (see e.g. [52, 83]).
- Similarly, a popular method to derive **driver drowsiness** is also based on image recognition. In order to calculate the PERCLOS (percentage of eye closures) score, the current state of the driver's eyes (open/closed) has to be assessed. The reason to exclude this feature is the same as for lane deviation: an implementation on the current hardware would decrease performance. Furthermore, the detection does not reliably work if the driver is wearing glasses.
- The ability to record sound was left due to the low priority of this feature. The **microphone** was mentioned in the interviews only twice, once in the context of stress levels for cyclists (P6) and the second time in regard to ethical concerns (P8). The decision to not include sound recordings was made based on two considerations: Firstly, bicycles are not the primarily targeted road users for the proposed system. Secondly, recording sound in the personal space of a car leads to serious privacy issues. Only few of the related projects made use of the microphone and never extensively: Bergasa et al. [42] leveraged the microphone

to detect the blinking sound as an additional indicator for correct lane changes. In light of these reasons, the decision was made to not incorporate the microphone.

- In the version presented in this thesis, the optional connection to an **OBD-II adapter** is used to acquire accurate speed, engine speed and fuel consumption values. Even though this data adds significant value to the platform, these three values are only a small subset of the information stored in the vehicle's internal system. The OBD-II standard specifies a long list of commands, however, car manufacturers have no unified implementation of this protocol. In practice, this means that a command will yield valid data in one car model but not in another. As a result, only the most common commands were implemented in this project. Additional commands of interest such as the ones for pedal usage or steering wheel angle would have to be added dependent on the specific car model used. A possible solution would be to integrate an additional step into the setup process in which researchers can manually input the specific command codes they are interested in. All in all, only three of the standard OBD-II values are ready to be used whereas more specific commands could be added at a later stage.
- The **event detection algorithms** implemented in the prototype are based on simple averaging filters and thresholds. On the one hand, this allows the system to run in real-time; on the other hand, the algorithms are neither as flexible nor as accurate as the more sophisticated approaches discussed in 3.2. Especially the area of machine learning (see e.g. [7, 50]) promises to deliver event detection that is tailored to different driving scenarios.

6.2 System Design

Before the actual implementation, the key states and interactions of the application were designed. Figure 6.1 depicts a state diagram detailing the process from the system's first start to the end of a study drive. After the initial study setup, the application transitions into a waiting behaviour, applying a hybrid trip detection algorithm. Once a significant vehicle movement is detected, the system changes into the data collection state where all specified sensors are active. As soon as the end of the trip is detected, the questionnaire is being shown to the user (if activated). After that, the system transitions back into the waiting state. Out of these states, the following paragraphs focus on the solutions designed for study setup, questionnaires and trip detection.

Study Setup: One of the key actions a researcher has to perform within the application is to set up a new study. To satisfy NFR 2, this action was divided into several smaller steps that guide the researcher through the process.

1. **Study Details**: In the first step of the study setup, the details about the research study and the participant can be entered. This includes a study identifier and name, a unique participant identifier and the participant's gender and age.
2. **Sensors**: The next screen allows specifying which of the core sensors are active during data collection. A researcher can independently activate or deactivate individual sensors out of that list.
3. **OBD-II Adapter**: This step allows connecting to an external OBD-II adapter. If this option is selected, the application will perform the connection setup and inform the researcher once the devices are connected.
4. **User Phone**: On this screen, the user's smartphone can be connected in order to log incoming notifications and screen touches. The procedure requires the researcher to start a small application on the user's phone to authorise the Bluetooth connection.
5. **General Settings**: Here, detailed settings of the data collection process can be adjusted. This includes the option to log data and define thresholds, sampling rates and video resolutions.

6. **Camera Calibration**: Once the study is set up, the application prompts the user to mount the smartphone on the windshield. If the cameras are activated, a preview of the video stream is shown to help adjusting the phone's position.

However, since the use of external devices is optional, researchers have the option to skip step three and four.

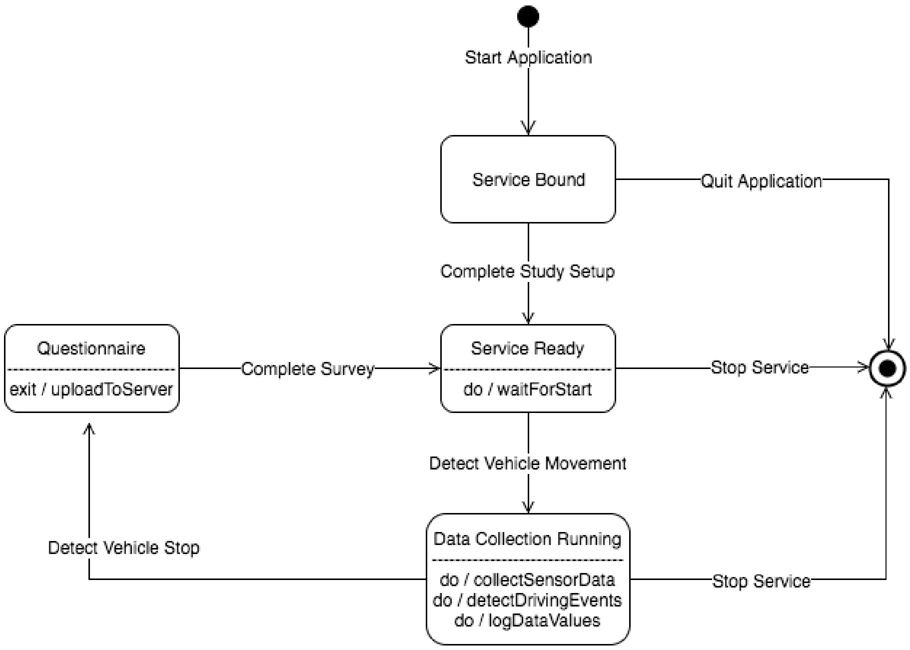


Figure 6.1: State Diagram of the data acquisition application.

The design of the above mentioned six-step setup was informed by paper sketches. During the process, the envisioned application interface transformed from a simple mockup with only two screens into the more stretched out procedure that was finally implemented (see Figure 6.2).

Start- and Stop-Detection: FR 2 requires the application to automatically detect the start and end of a trip. To achieve this, several different implementation options were identified. In previous research projects, the start of a trip was either detected by monitoring the longitudinal acceleration of the smartphone (e.g. [43]) or the speed measured by GPS signals (e.g. [42]). For the application developed in this thesis, the decision was made to combine both approaches into a hybrid trip start detection. The system checks the phone's location every five seconds and triggers a trip start if it is significantly different to the last measured location (Δd > 15 meter). Simultaneously to this procedure, a significant motion sensor is monitoring the accelerometer data, triggering a start after detecting "a motion that might lead to a change in the user's location; for example walking, biking, or sitting in a moving car"[14]. Since the phone is assumed to be mounted in the vehicle, any motion detected by this sensor must be a movement of the vehicle. This hybrid approach combines the suggestions found in literature and promises to accurately detect the beginning of vehicle movement. Similar to the start detection, the trip end detection also applies a hybrid method: a trip ends when either the OBD-II adapter reports an engine speed and speed value of zero or when the GPS calculated speed is close to zero for a period of 45 seconds.

Questionnaires: The surveys are conceptually tightly coupled with the end of a trip. Once the stop-detection algorithms detect the end of a drive, the application can display a predefined questionnaire. In the process of designing this interaction, several different use cases were explored.

[14]https://developer.android.com/guide/topics/sensors/sensors_motion.html, Accessed 13.09.2016

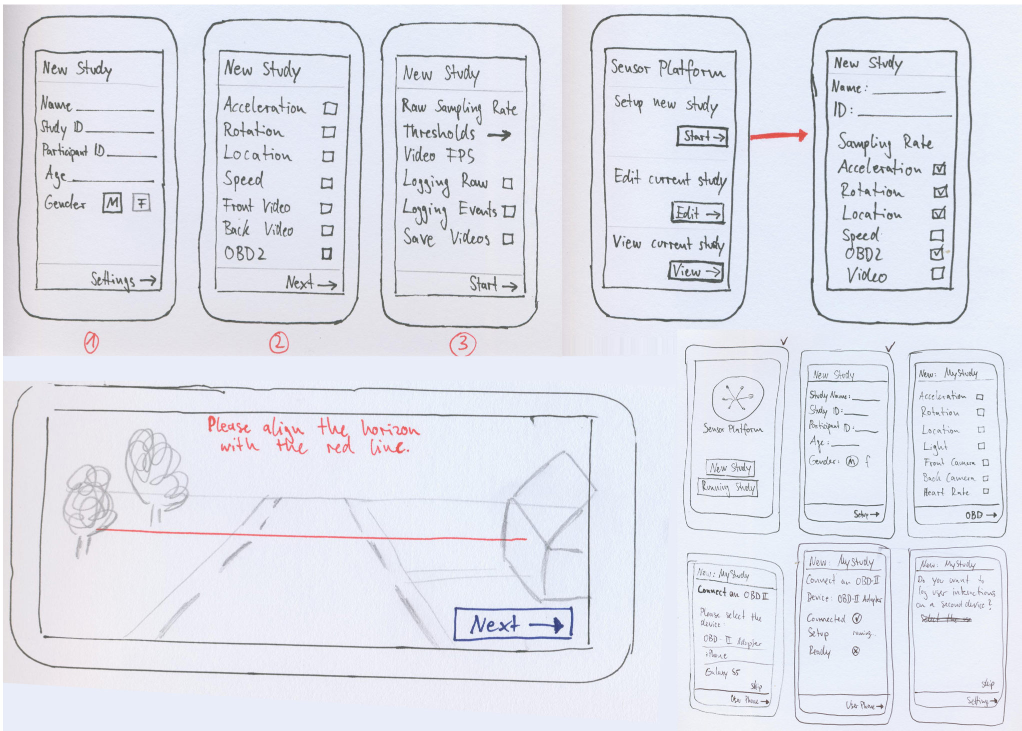


Figure 6.2: Sketches helped to create and refine the steps and user interface of the study setup.

Among those, the key difference is whether the participant's personal phone is connected to the research phone. If both phones are connected, the survey could either be displayed on the research phone, or, more conveniently, on the participant's personal phone.

Another use case includes the use of Near Field Communication (NFC) tags attached to the mount in order to detect the removal of the research phone (compare Figure 6.3). The simplest way to design the interaction is to show the survey on the mounted phone immediately after a trip end is detected and notify the participant by turning on the screen and playing a short sound. For the first development iteration, this option was chosen to be implemented.

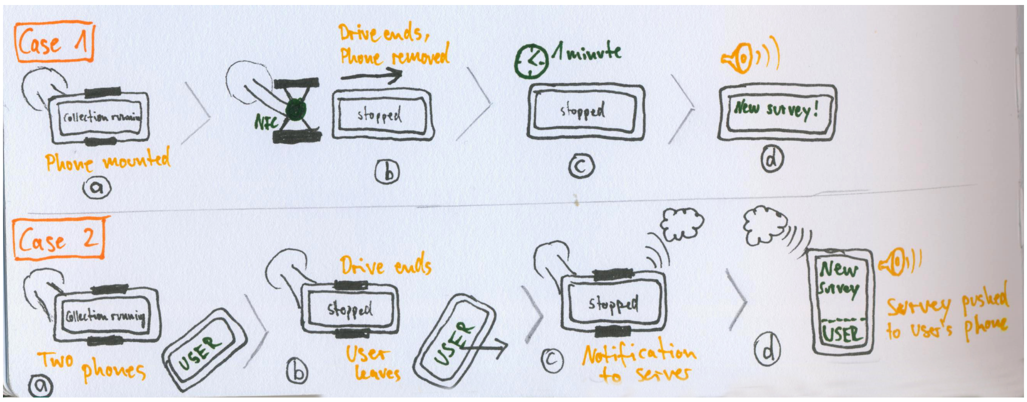


Figure 6.3: Several use cases for questionnaire interactions were explored.

6.3 Hardware

In order to fully operate *Sensor Platform*, the following four devices are needed:

1. the main smartphone (required),

2. a second smartphone (optional),
3. a fitness wearable that can measure heart rate (optional), and
4. an OBD-II adapter (optional).

Only the main smartphone is required to use the core functionality. The use of the other three devices is optional. Figure 6.4 shows the devices used during development.

Main Smartphone: A major part of the implementation was driven by the choice of the operating system (OS) powering the main smartphone. Three different mobile OS were examined as possible choices: Android, iOS and Windows Phone. Within the scope of this project, only one operating system could be targeted. Out of the three options, Android is the best fit for the proposed project: iOS and Windows Phone do not offer an easy way of distributing the app because they require uploading the application to the respective stores. Android applications on the other hand can be easily installed with just one downloaded file. This streamlines the setup process participants have to go through before being able to commence the study. Another reason to choose Android as the operating system is the high market share: in June 2016, 61.6% of all smartphones in Australia were running on Android[15]. As the main smartphone for data collection, a Google Nexus 6P running Android 7.0 was used. For *Sensor Platform* to function properly, Android 5.0 or greater is needed. The Nexus 6P is equipped with all required sensors (see FR1 in Section 5.2) and has enough processing power to perform the algorithms (Snapdragon 810, 3GB RAM).

Second Phone: In order to test the connection to a second phone, the helper application was installed on a Samsung Galaxy S4 with Android 4.3. The minimum requirement for the second phone is Android 4.1.

Wearable: Besides the two phones, a Motorola Moto 360 (2nd Generation, Android Wear 1.5) showcases the ability to gather heart rate data. The Android Wear ecosystem was chosen because of the good integration with the Android OS running on the main smartphone. Alternatives such as FitBit do not allow such a tight coupling between the devices and require the use of an online API.

OBD-II: Lastly, an ELM327 mini, manufactured by ELM Electronics[16] is used in the implementation to access the vehicle data. It provides a Bluetooth interface which allows connecting and communicate between the main phone and the ELM327. In order to bypass the need to connect the ELM327 to a real vehicle during development, an OBD-II emulator manufactured by Freematics[17] provided a customisable simulation of vehicle data.

(a) Google Nexus 6P (b) Motorola Moto 360 2nd Generation (c) ELM 327 mini OBD2 Adapter

Figure 6.4: The hardware components used for the prototype implementation.

[15] http://www.kantarworldpanel.com/global/smartphone-os-market-share/, Accessed 17.08.2016
[16] https://www.elmelectronics.com/ic/elm327/, Accessed 06.09.2016
[17] http://freematics.com/store/index.php?route=product/product&product_id=71, Accessed 08.09.2016

6.4 Implementation

The research platform consists of two major parts: a smartphone application running the data acquisition service and a web server handling data storage and information visualisation. It should be noted that the focus of the development phase lay on the Android application. In the next two subsections, the implementation details of these two parts are described.

6.4.1 Android Application

The implementation of the main Android application was done in Android Studio 2.2. The code base is structured into three components (see Figure 6.5), namely the **user interface**, the **core functionality** and procedures for connecting to and communicating with **external sensors**. In the following paragraphs, the three components are explained in detail.

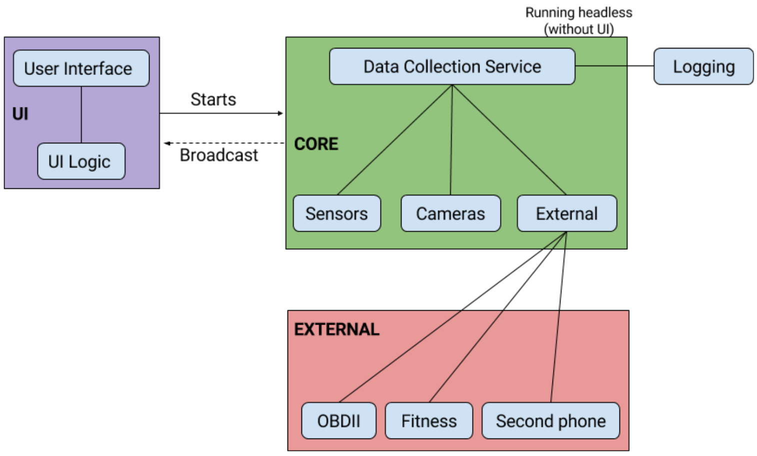


Figure 6.5: Application structure: user interface, core functionality and external components.

User Interface
The user interface consists of only a small number of screens because the central part of the application – the data collection – is running as a service in the background. Any interaction that requires the participant to control the application during a trip is a safety hazard as it draws the driver's attention away from the road. Consequently, the system was designed to require as little user interaction as possible after the initial setup.

The initial creation of a new study on the phone is the main interaction performed by the researcher. In order to satisfy FR 7 and NFR 2, this operation was divided in five steps as described in Section 6.2. The main aim is to guide the researcher through the creation process. Following Shneiderman's *Eight Golden Rules of Interface Design* [84], each screen highlights the current step of the process and gives hints to inform the researcher's decision (see Figure 6.6). In steps three and four, the user interface provides feedback on the current connection status, notifying the user when the Bluetooth connection was successfully established.

During data collection, the screen is meant to be completely turned off. Conceptually, the driver should not be influenced by the phone's presence as this might alter the driving behaviour. However, if a participant or researcher wants to get feedback on the application's operation, the data collection screen can either show a two-state display ("Waiting for trip to start." / "Data collection running.") or a detailed live view of all values and events (compare Figure 6.7). The latter is not designed to be used while driving but can offer insights during studies where an expert observer is present in the car. Furthermore, the live view might enable future work that features safety interventions, e.g. *CoastMaster* [9].

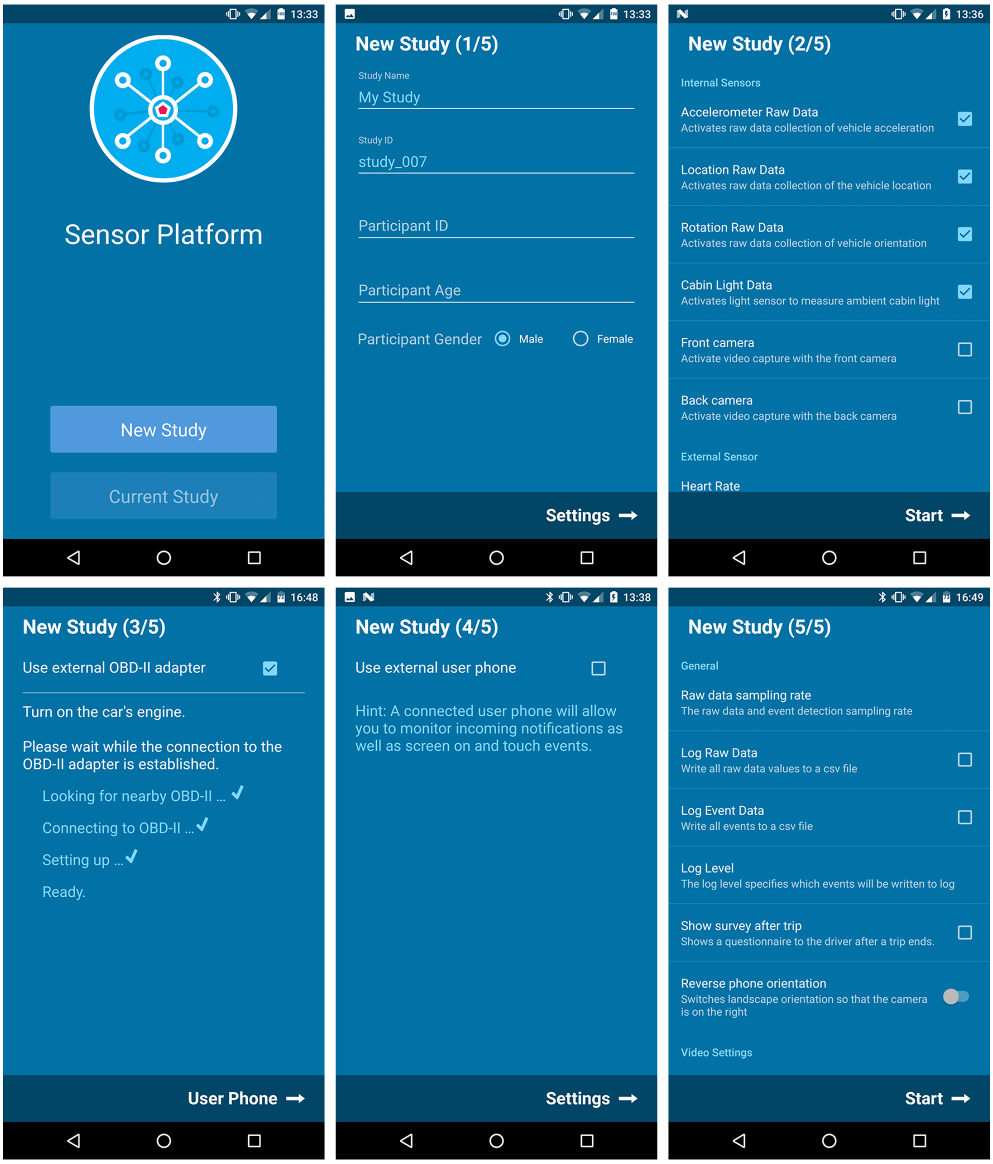


Figure 6.6: Creating a new study involves six steps, of which the first five are depicted here. In the last step, the cameras are calibrated.

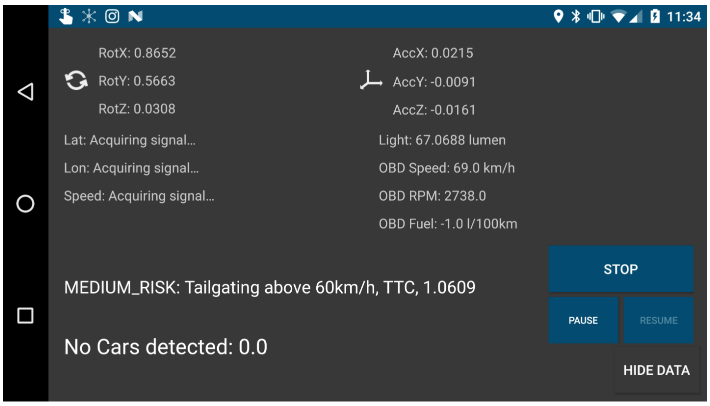


Figure 6.7: The live view of all collected data values and detected events is meant to be used by expert observers but not the driver.

Once the end of a trip is detected, the system can automatically display the questionnaire. As the survey is meant to be answered immediately after the drive, the screen will turn on in order to make the user aware of the new task. The questionnaire interface consists of the current question and five radio buttons representing a 5-point Lickert scale (compare Figure 6.8). After the last question is answered, the application transitions back to the data collection screen.

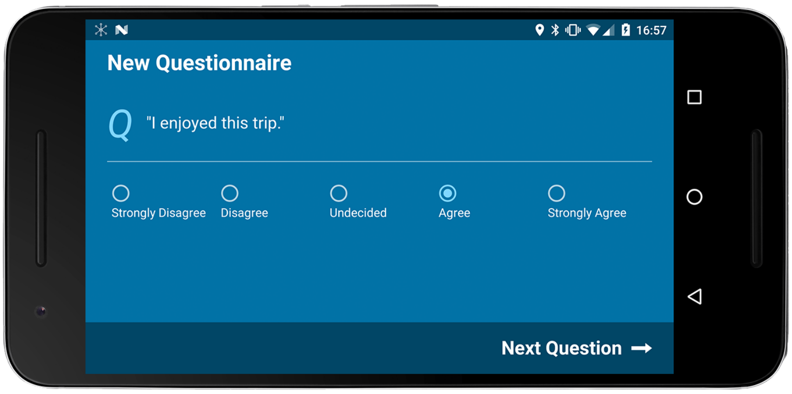


Figure 6.8: Questionnaires are displayed after a trip ends.

Core Functionality

At the core of *Sensor Platform* runs the data collection service, called `SensorPlatformService`. After a new study is set up by a researcher, this background service is started. The service runs headless; it does not need the application to be visible to function correctly. With this method, data collection can not be stopped by accident and is not restricting the phone's operation (except for a decrease in performance).

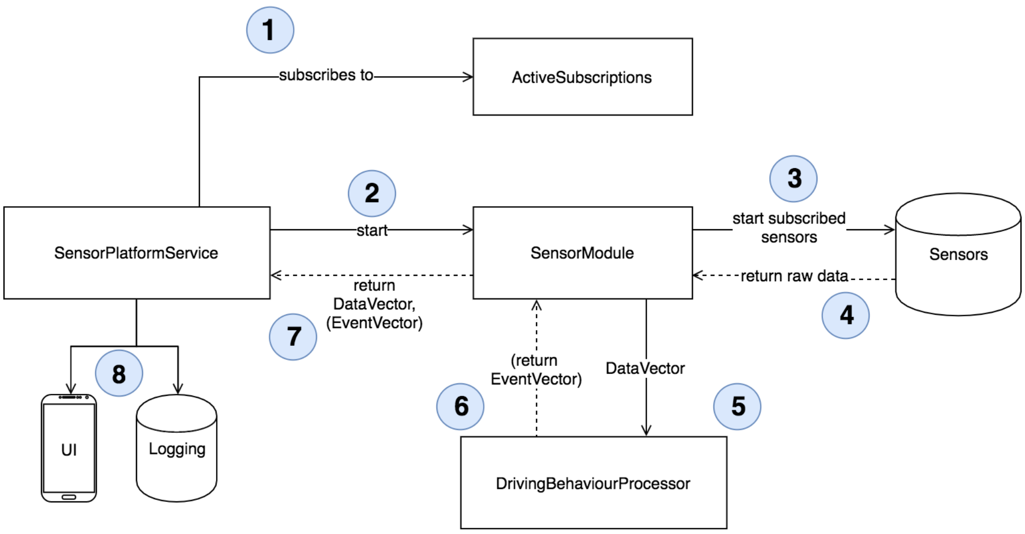


Figure 6.9: The eight steps of *Sensor Platform's* data collection.

Figure 6.9 shows the basic steps repeated during data collection: The `SensorPlatformService` subscribes to the sensors, both internal and external, specified by the researcher (1) and instructs the `SensorModule` to start the data collection (2,3). The `SensorModule` class gathers the incoming sensor data for a specified time interval (e.g. 250ms) before combining it into a data structure called `DataVector` (4). This `DataVector` is then piped into the `DrivingBehaviourProcessor` (5), where the event detection is performed. This processor has the capability to detect hard accelerations, hard braking, sharp turns and speeding. If the algorithms detect a critical driving event, an `EventVector` is instantiated and sent back to the service's main sensor hub, the `SensorModule` (6). All instances of `DataVector` and `EventVector` are handed back to the `SensorPlatformService` (7), which then performs the operations for logging and broadcasting the data (8). The user interface can listen for these broadcasts and display the so received data in real-time.

Filtering Algorithms
Raw data collected from the sensors contains a significant amount of noise, in particular the accelerometer readings. Therefore, in steps (4) and (5), the raw data values are filtered to remove noise artefacts and smoothen out sudden spikes in the data. Starting from the unprocessed accelerometer values, different filtering algorithms were explored, most prominently among them the **exponential moving average** and a one-dimensional **Kalman Filter**. Both algorithms are fit to smoothen a series of data values, however, they differ in the details.

The **exponential moving average (EMA)** calculates the average of a sliding window by applying weight factors that decrease exponentially. As such, the EMA reacts faster to most recent data changes than a simple moving average. It incorporates a coefficient α that represent the degree of weighting decrease, i.e., it controls the amount of smoothing. In many applications, α is based on the number of samples and defined as:

$$\alpha = \frac{2}{N+1} \tag{2}$$

The EMA is very fast to compute because each iteration only needs the previous EMA value and the most recent raw value as input. The calculation of the EMA is as follows:

```
private double prevEMA;
private double alpha = 0.35; // or calculate based on sample size

public double getEMA(double currentValue) {
    if(prevEMA == null) { // first step
        prevEMA = currentValue;
        return prevEMA;
    }

    double result = alpha*currentValue + (1-alpha)*prevEMA;
    return result;
}
```

The **Kalman Filter** implemented for filtering acceleration values is a one-dimensional filter. As such, it differs from the usual application of an extended Kalman Filter. As introduced by Welch and Bishop [13], the algorithm usually combines multiple data sources for an accurate prediction. In the specific case of this project, only one data source is fed into the filter. However, the main advantage of dynamically adapting parameters still remains in comparison to the EMA. Based on the initial parameters for process and measurement noise, the algorithm first predicts a new value and then verifies it against the real measurement. The following implementation of an one-dimensional Kalman Filter is used in this project:

```
private double q = 0.2; // process noise covariance
private double r = 3; // measurement noise covariance (confidence
    in sensor)
private double p = 1023; // estimation error covariance
private double accZ = 0; // initial value

private double calculateKalmanValue(double accZ_measured) {
    // prediction step
```

```
        p = p + q;

        // measurement update
        double k = p / (p+r);
        accZ = accZ + k * (accZ_measured - accZ);
        p = (1-k) * p;

        return accZ;
    }
```

Both filtering methods are implemented in the prototype with the Kalman Filter being the default.

Image Processing

The image processing pipeline is separated from the other sensors. While the `SensorModule` is responsible for collecting data from multiple sensors such as the accelerometer and GPS, the `ImageModule's` sole task is the operation of the front and the back camera. However, the general sequence of data collection is similar: The `ImageModule` first acquires raw image data before calling the `ImageProcessor` to extract the required event information. In this stage, the front camera image is processed to detect driver distraction whereas the back camera image is used to monitor following distance. Because image processing in Java is too slow for real-time analysis, this application uses the Android NDK[18] in combination with OpenCV 3.0[19] to perform algorithms written in C++ and C, namely the face detection and front vehicle detection. For these two detection procedures, the raw images are first converted to grayscale images, then resized to a resolution of 320x240 pixels.

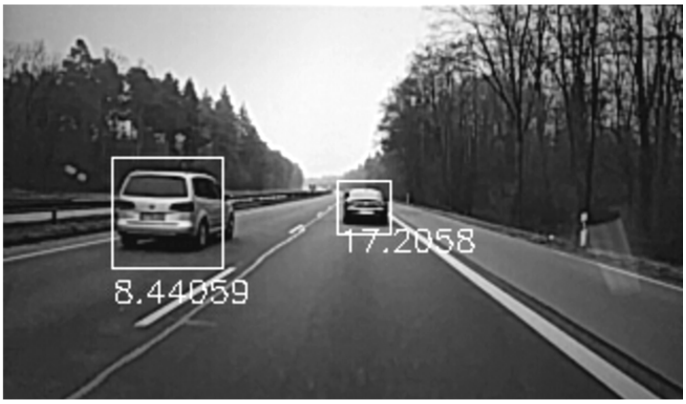


Figure 6.10: The application detects rear-ended vehicles and estimates the following distance based on the average vehicle width.

After that step, a cascade classification is performed to detect the driver's face and vehicles in front. The used cascade source files were developed in a previous project at QUT. This classification step produces a set of bounding rectangles around each detected object. In the case of car detection, a region of interest (ROI) is defined prior to the cascade classification which reduces the processing time for each frame and limits the number of false positives. The need for such a ROI became clear while testing the algorithm as it often detected cars that were not in the same lane (compare Figure 6.10). In order to then calculate the following distance, the formula described in Section 3.2.2 is applied. The width of each bounding rectangle and the average width of a standard car are used as input. Figure 6.10 shows the result of one such car detection and distance calculation step.

[18] https://developer.android.com/ndk/index.html, Accessed 06.09.2016

[19] http://opencv.org/, Accessed 06.09.2016

Speeding Detection
Another challenge was to detect overspeeding: While accelerations, braking and turns can be extracted with only sensor data as input, speeding events rely on map data from an external service. On basis of the recorded GPS signals, a reverse geocoding algorithm has to be performed. Reverse geocoding is the process of converting a geolocation to a readable address or place name [85]. This information is required to query information about speed limits. For speed limit data, several services were compared:

- *Google Maps Roads API*[20]: Google Maps is the most popular map service currently available and offers an API to query speed limits based on a set of geo locations. However, at the time of writing, the service requires a commercial "Google Maps API for Work" subscription. As a result, the Google Maps Roads API is not suitable for an open-access research tool.
- *Wikispeedia*[21]: This platform is a crowdsourced open access database of speed limits. Due to a missing documentation and no information about how active the community is, this option was ruled out as well.
- *OpenStreetMap Overpass API*[22]: OpenStreetMap is a well known open-source alternative map service. It is based on a crowdsourcing approach, allowing anyone to enter or edit map data. The Overpass API offers access to any information contained within the OpenStreetMap databases and is thoroughly documented. The only drawback of this service is that information on speed limits is not available for all roads. Nonetheless, the Overpass API was deemed the most suitable option for this project.

Similarly, in order to include information about the current weather condition, an external web service is necessary. The application queries the *Yahoo Weather API*[23] every 300 seconds, acquiring information about temperature, wind conditions and atmospheric values (humidity, pressure).

Data Storage
The collection and fusion of the sensor data is only the first step. In order to allow researchers to analyse the data, log files of all collected values must be stored. In the case of the implemented system, logging is done in the form of CSV files. For each trip, the system creates two new files, uniquely identifiable by a `tripID` generated by combining the `participantID` and a `timestamp` of the trip start. Each line of the raw data CSV file consists of the information stored in one `DataVector`. The event data file contains all the detected events of the trip, along with a description, two values and the filenames of the respective video files. Each event is linked to one entry in the raw data file, identified by the primary keys `tripID` and `timestamp`. That way, the data can be processed further, e.g. it can be imported into a database or analysed with statistics software.

Videos and log files are stored on the phone's flash memory, potentially allowing users to access, review and edit data. Giving a participant access to the data has a clear benefit: When conducting a naturalistic driving study, ethical considerations play a major role. Recordings of illegal driver behaviour can lead to legally difficult situations. As one expert pointed out during the requirements analysis, it is unclear if the principal researcher is obliged to report an incident to the police if evidence is present in the data. Allowing participants to review the data prior to uploading or analysis can lighten that ethical responsibility for the researchers. Of course, this procedure risks reducing the validity of results due to potential removal of critical events. For that reason, this specific feature was not directly integrated into the application's user interface but could be implemented later as an add-on.

[20]https://developers.google.com/maps/documentation/roads/speed-limits, Accessed 07.09.2016
[21]http://www.wikispeedia.org/, Accessed 07.09.2016
[22]http://overpass-api.de/, Accessed 07.09.2016
[23]https://developer.yahoo.com/weather/, Accessed 07.09.2016

External Sensors
Optional external components as required by FR 4 are the OBD-II adapter, a second smartphone and a fitness wearable. The connection process is performed during the study setup. In the case of a broken or lost connection, the system tries to automatically reestablish the connection. If activated, the system is programmed to query the **OBD-II adapter** at a regular interval. For communicating with the adapter and parsing the byte values of each response, the open-source library *obd-java-api*[24] is used.

In order to log user interactions and notifications on a **second smartphone**, a helper application has to be installed on the second phone prior to the start of the study. This small application allows setting up a Bluetooth connection between the user's and the research smartphone. Once started, a background service detects the events for "screen-on", "screen-touch" and "incoming notification" and sends a Bluetooth message to the connected research phone. Additionally, the logged event specifies which application was running in the foreground at the time of interaction. From a road safety perspective, it is a considerable difference whether the driver is interaction with a navigation app or a social media software. Combined with the number of screen touches, this information allows researchers to judge the severity of the distraction.

Lastly, *Sensor Platform* implements a connection to **Android Wear smartwatches** in order to sense the driver's heartbeat. As a prerequisite, the smartwatch has to be registered as a connected watch by the Android OS. After this, the pairing of phone and watch is handled by the operating system. While sensing the heartbeat, the system outputs the heart rate reading and its accuracy. Consequently, values with unreliable accuracy are ignored by the algorithm and not reported to the research phone.

In summary, the developed main application can communicate with three external sensing devices to broaden the data collection abilities, thus fulfilling the requirement identified in Section 5.2.

6.4.2 Web Server

Complementary to the Android application, a web server consisting of a database backend and a front-end was implemented. The web server's main purpose is to provide access to the database containing raw and event data and to serve a simple frontend that shows and visualises the data. In the scope of this project, only a minimal proof-of-concept web server and interface were implemented. While the focus of this thesis is on the data acquisition system itself, the web server showcases a possible direction of future work, i.e., research on how the collected data can be visualised, e.g., by highlighting events and generating statistics across several drives.

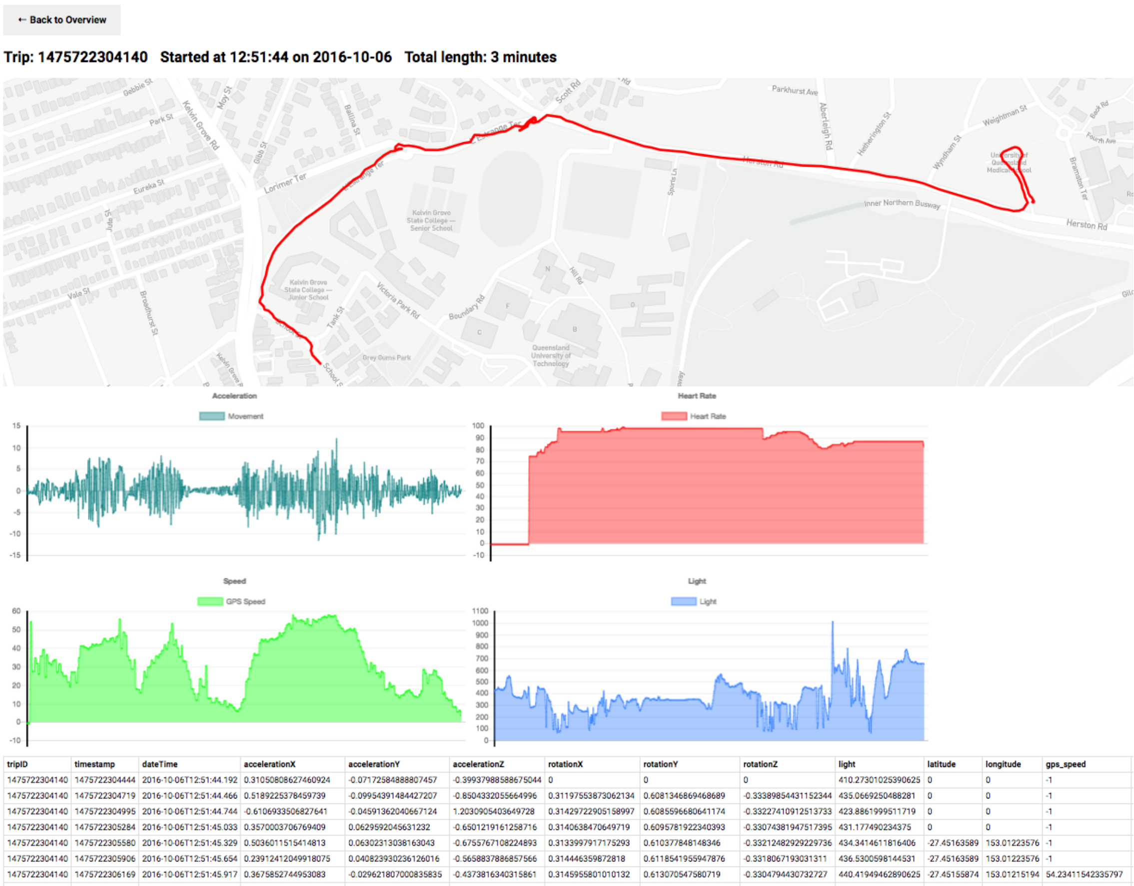
Running on *NodeJS 5.0* and *ExpressJS 4.14*, the web server connects to a non-relational MongoDB database. Access to the data is offered through a simple REST API. Clients can save data to the database by posting the CSV files generated by the Android application described in the previous section. The server then automatically parses the incoming data and saves it both as file and as a document in the database.

The web frontend lists all stored trips, identified by the `tripID` as unique key. The user can select an individual trip to show detailed information. This detailed view includes a map plotting the trajectory, several line charts to visualize the data (accelerations, speed, heart rate, cabin light) as well as a table containing the raw data itself. Figure 6.11 shows the web interface that allows exploring a single trip.

To summarise, this chapter explored the **second research question RQ2** and presented the steps of designing the system before implementing a first version of the research platform. The main focus of the development process lay on the Android application for data collection. Based on the

[24]https://github.com/pires/obd-java-api, Accessed 20.09.2016

requirements specification, the system's architecture and interaction flow was prototyped using paper sketches and UML modelling. Following the interaction design model, an interactive and functional version of the system was implemented. Due to the limited scope of this thesis, some features were identified as possible additions for future versions. Besides the Android application, the chapter also touched on the concept and implementation of a simple web server that can store and visualize the collected driving data.



tripID	timestamp	dateTime	accelerationX	accelerationY	accelerationZ	rotationX	rotationY	rotationZ	light	latitude	longitude	gps_speed
1475722304140	1475722304444	2016-10-06T12:51:44.192	0.31050808627460924	-0.07172584888807457	-0.39937988588675044	0	0	0	410.27301025390625	0	0	-1
1475722304140	1475722304719	2016-10-06T12:51:44.466	0.5189225378459739	-0.09954391484427207	-0.8504332055664996	0.31197553873062134	0.6081346869468689	-0.33389854431152344	435.0669250488281	0	0	-1
1475722304140	1475722304995	2016-10-06T12:51:44.744	-0.6106933506827641	-0.04591362040667124	1.2030905403649728	0.3142972290515899 7	0.6085596680641174	-0.33227410912513733	423.8861999511719	0	0	-1
1475722304140	1475722305284	2016-10-06T12:51:45.033	0.3570003706769409	0.0629592045631232	-0.6501219161258716	0.3140638470649719	0.6095781922340393	-0.33074381947517395	431.177490234375	0	0	-1
1475722304140	1475722305580	2016-10-06T12:51:45.329	0.5036011515414813	0.06302313038163043	-0.6755767108224893	0.3133997917175293	0.610377848148346	-0.33212482929229736	434.3414611816406	-27.45163589	153.01223576	-1
1475722304140	1475722305906	2016-10-06T12:51:45.654	0.23912412049918075	0.040823930236126016	-0.5658837886857566	0.314446359872818	0.6118541955947876	-0.3318067193031311	436.5300598144531	-27.45163589	153.01223576	-1
1475722304140	1475722306169	2016-10-06T12:51:45.917	0.3675852744953083	-0.029621807000835835	-0.4373816340315861	0.3145955801010132	0.613070547580719	-0.3304794430732727	440.41949462890625	-27.45155874	153.01215194	54.23411542335797

Figure 6.11: A web-frontend highlights how researchers can visually explore the data.

7 Evaluation

In order to assess the overall quality, usability and accuracy of *Sensor Platform*, an evaluation through qualitative and quantitative methods was conducted. Following the thesis' user-centered approach, the aim was not only to test the application from a technical point of view but also to gather feedback from the later users, i.e., road safety experts. Furthermore, the evaluation promises to highlight the (dis)advantages of the explored low-cost approach.

7.1 Objectives

The evaluation process was therefore divided into two main blocks: the first block consists of a focus group and a user study with road safety experts; for the second block, several test drives were performed to judge the platform's data collection abilities. Each of the two evaluation phases aims to answer one of previously introduced research questions.

RQ3a: *How do road safety experts judge the low-cost system in terms of usability (i.e. effectiveness, efficiency and satisfaction)?*

Chapter 5 specified that researchers without technical knowledge or previous experience should be able to operate the platform without errors or support (NFR 2). The previous chapter explained the steps taken towards creating a coherent and easy to use procedure for setting up a new study. As this requirement is crucial for a successful deployment of the application, this question is a specific focus of the evaluation. Furthermore, the exploration of this question promises to identify areas of improvement, a valuable source of information for future versions of the system.

RQ3b: *How does the smartphone-based system perform in terms of accuracy?*

The application's main purpose is to collect naturalistic driving data. As such, the accuracy of the collected data, including critical driving events, must satisfy the needs of road safety researchers. While studies may differ in regard to the acceptable margin for error, it is safe to assume that every research study aims to get as accurate data as possible. An evaluation of this question will identify the strengths and limitations of the low-cost approach and will therefore help road safety experts to judge whether the application is suitable for a specific driving study.

This chapter reports on the two studies that were designed to answer the above questions and provides details on methods, procedures, and results.

7.2 Focus Group

Prior to the two studies, a focus group gave a small number of experts the opportunity to provide preliminary feedback on *Sensor Platform*. Furthermore, the study design for an assessment of the accuracy was discussed.

7.2.1 Procedure

Participants (*N*=4, 3m, 1f) were recruited among the experts who were interviewed earlier in the project. The focus group started with a presentation of the requirements specification process and a live demonstration of the application's features and interface. The goal of this first phase was to allow participants to get a good understanding of the capabilities and limitations of the developed system. Before continuing with the procedure, the experts had the chance to ask questions about *Sensor Platform*. In the next step, participants were asked to provide feedback. Specifically, they were instructed to focus on possible use cases, missing features and suggestions for improvement.

To conclude the focus group, participants were asked for advice on how to design the technical evaluation of the platform. In particular, they were encouraged to think about what form of evaluation would help road safety experts to judge the suitability of the research platform for future studies.

7.2.2 Results and Discussion

The following paragraphs summarise and discuss the results of the focus group, specifically the benefits of *Sensor Platform* as mentioned by experts, possible future functionalities, and considerations regarding the technical evaluation.

Benefits
Overall, participants provided positive feedback. One expert stated that he is considering to use *Sensor Platform* as a data collection tool in his research, because, as he said, “it is **much easier to deploy**” (P1) than traditional data acquisition systems. It would allow collecting driving data of a large number of people without facing high costs. P4 was interested in **physiological data** and therefore liked the option to connect a fitness wearable to *Sensor Platform*. The participant inquired how difficult it is to support more advanced physiological measuring devices which shows the scepticism towards the accuracy of consumer electronics in that area. Furthermore, the phone’s internet connection was seen as a big advantage, as participants were thinking about “how to link in **other sources of data**” (P3). By supporting requests to external APIs, the driving data can be enriched with information about the environment. While speed limits and the current weather are already implemented in the current version, more specialised APIs, e.g. the location of roadworks, could be added if required for a specific study.

Functionalities
Besides the highlighted advantages, the participating experts were interested in additional functionality. P1 proposed to implement a way to **remotely set up** the application in order to reduce the effort of deployment even further. Specifically, this means that the responsible researcher can remotely manipulate application settings on a participant’s phone by using a central command server. This has two advantages: firstly, it would would make it possible to bypass the setup process almost entirely. A participant would only need instructions on how to mount the phone in his vehicle; after that, the researcher could start the study by sending a command to the phone over an internet connection. Secondly, a remote connection would allow changing settings during an active study, e.g. adapting the video resolution in order to save storage space.

Participants would have liked more advanced image recognition functionalities such as **monitoring pedestrians and cyclists** (P2), or **drowsiness and lane weaving detection** (P4). While such functionalities have been implemented using image processing in professional dedicated hardware such as *MobilEye Shield+*[25], or in specialised smartphone applications (e.g. [6, 52]), it is challenging to implement these algorithms as part of a real-time data collection running on a mobile device. These features represent a valuable addition but are not part of the core functionality of current data acquisition systems. The reasons not to focus on drowsiness and lane weaving detection were already discussed in Section 6.1 whereas pedestrian and cyclist detection was not a topic of the earlier interviews. The experts’ interest in such features implies that they were satisfied with the existing core functionality but wanted to learn more about possible future directions.

Considerations for Evaluation
P1 made clear that **data accuracy** is important to him: an evaluation of that aspect would be necessary “to see how much you can **trust** it”. According to participants, comparing *Sensor Platform* to an existing professional data acquisition system is a good way to assess its accuracy. Since it can be expected that the two compared systems do not have the exact same functional

[25] http://www.mobileye.com/shieldplus/, Accessed 28.11.2016

range, only certain data values stand out for comparison. For that reason, P3 suggested to group the various features into several categories. A "core-functionality" category would include the data collection capabilities that are usually part of data acquisition systems whereas other features such as logging the user phone interactions could be flagged as "experimental" and evaluated separately. Furthermore, P3 also pointed out that some aspects of data collection on a smartphone are already known to have certain limitations, e.g. the GPS signal. As a result, an evaluation of *Sensor Platform* would not have to include measurements like the **GPS accuracy**. Yet, despite the known limitations of GPS sensors in mobile phones, it still makes sense to compare that type of data to the professional system because some important metrics are based on the GPS signal, most importantly the vehicle speed. An evaluation of the GPS signal can show how much the speed calculations are influenced by erroneous location updates.

Besides that, P1 highlighted his interest in operational information that goes beyond the actual collected sensor data: **storage requirements and battery consumption**. Consequently, the results of the technical evaluation should include such information. As this information is not automatically stored in the log files, P1 raised a valid point, especially with regard to the limited storage space that most mobile phones offer. Also, a high battery consumption would imply that the phone needs to be constantly connected to a power source which complicates the installation in the car.

In addition to that, experts pointed out that it would be helpful to have a list of recommended smartphone models. That way, researchers do not have to compare different models and can be ensured of the full feature set.

7.3 In-Vehicle Expert User Study

The second part of the expert evaluation was designed to provide both qualitative and quantitative feedback on the usability of *Sensor Platform*. Usability is not a quality that can be measured in an absolute sense; however, as described in ISO 9241-11 [86], it encompasses multiple criteria such as effectiveness, efficiency, and satisfaction, which can be more easily assessed. For that purpose, a user study with N=10 participants was conducted. The focus of the study lay on the setup procedure of the research platform, i.e. the process that a road safety researcher has to perform in order to prepare a vehicle for a NDS. This involves setting up the smartphone app, connecting external sensors and installing the phone in the vehicle.

7.3.1 Participants

N=10 Participants (7m, 3f, M_{Age} = 34.5, SD_{Age} = 9.12) were recruited among road safety researchers and interaction designers. In case a participant was not familiar with certain road safety terminology, a fact sheet explaining relevant terms was handed out before the start of the study.

7.3.2 Procedure

The study took place on a parking lot next to the study vehicle (see Figure 7.1a). Before the start of the study, each expert was handed a short summary about *Sensor Platform*, including a description of the data collection abilities, event detection, and available optional sensors (see Appendix B). Once the expert acknowledged having read and understood the summary, the principle researcher explained the scenario for the study. For that, experts were asked to imagine that they are currently conducting a naturalistic driving study using *Sensor Platform* as the data acquisition system. Setting up the application and installing it in the vehicle was presented as the overarching task for the in-vehicle evaluation. In order to make the procedure comparable, the main task was divided into six smaller tasks for each of which the principal researcher handed the expert a card with instructions (see Appendix C). Furthermore, a box containing all of the required equipment (research phone, user phone, OBD-II adapter, windshield mount) was given to the expert (see Figure 7.1b).

(a) Study car in which experts had to install the DAS.

(b) Box of equipment: research phone, user phone, OBD-II adapter, windshield mount.

Figure 7.1: The setup of the study.

While performing the tasks, experts were instructed to think out loud to explain what they are seeing on the screen and which steps they are taking in order to achieve their goal. The principal researcher took notes of all comments. For participants without prior experience with OBD-II adapters, the principal researcher explained its purpose and demonstrated where to connect it to the vehicle.

After completing the six subtasks, each participant had the chance to provide additional qualitative feedback in a short semi-structured interview. Participants were first asked to describe their general impression of the setup procedure, followed by questions that further explored the mentioned issues.

Lastly, the participants were asked to fill out the *System Usability Scale (SUS)*. The SUS is a short survey consisting of ten Lickert scale items and is widely recognized as a reliable tool for measuring the usability of a system [87]. Furthermore, it can be used accurately evaluate a system even with relatively small sampling sizes (N=8–12) [88]. The SUS questionnaire used for the study can be found in Appendix D.

The whole procedure lasted between 25 and 40 minutes, depending on how detailed the participants were describing their thoughts during the study.

7.3.3 Results and Discussion

The *System Usability Scale* yields a single number representing the composite measure of the overall usability of the application [87]. In order to calculate the final score, the Lickert scale items first have to be preprocessed before summing them up. For odd items, the score contribution is the scale position minus 1; for even items, it is 5 minus the scale position. The sum of these scores is then multiplied by 2.5, yielding a final SUS score between 0 and 100. It should be noted that the final score can not be interpreted as a percentage. Furthermore, SUS is not diagnostic, i.e., it does not specify which parts of a system cause trouble to a user.

Participants rated the usability of the *Sensor Platform's* study setup procedure with a score of M_{SUS}=81.0 (SD_{SUS}=10.08). In order to interpret this score, additional information must be taken into account. As Bangor et al. [89] describe, the SUS score is good for relative comparisons but hard to interpret as a single absolute value. Thus, Bangor et al. extended the questionnaire to match SUS scores against adjective ratings and a grade scale. As can be seen in Figure 7.2, a score of 81.0 positions *Sensor Platform* well within the "acceptable" range, marking its usability as "good" to "excellent". The good rating is also apparent when compared to the historical SUS average ($M_{SUS\ historic}$=69.69, $SD_{SUS\ historic}$=11.87) [90].

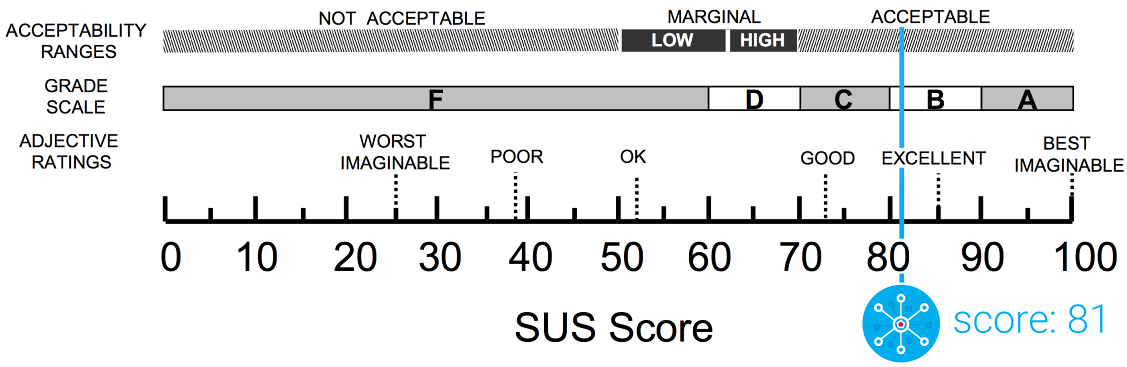


Figure 7.2: The scale proposed by Bangor et al. [89] helps to interpret a SUS score. Participants rated *Sensor Platform* with a score of 81, putting it between *good* and *excellent* and clearly inside the acceptable range.

The SUS result closely matches the participants' qualitative feedback during the study. The majority of experts found the setup process **easy to perform**, even on the first try. P6 found it "really simple", others called the procedure "straightforward" (P10). P4 described it as "well-designed" and was surprised how little time the setup process takes. The "long list of things to check" (P3) was experienced as a good level of detail that allows for customisation. One participant stated that the whole study setup feels similar to what researchers are used to: P9 said it "reminds me of a checklist" and performing the setup feels "like following a protocol".

However, answers indicate that this can be a disadvantage, too: One of the main findings is that **users tend to not read instructions** when they are presented in textual form. P8 felt the screens were too homogenous and would have liked more visuals during the instructions, because the "brain turns off" when everything is text-based and presented in a similar layout. Similarly, P1, P3 and P7 wished for more information on how to perform the specific operations, e.g. in the form of additional illustrations. They furthermore suggested to include a "help button" that can provide further explanations if needed.

These comments make clear that the application could benefit from two changes: firstly, **reducing the text length** of instructions and secondly, highlighting the required actions through **visualisations**. Of course, not all steps of the setup process are equally suitable for these changes. The more complex a task is, the more it benefits from a streamlined presentation.

A good example for this is the connection setup to a second phone, perhaps the most difficult task for participants. The step requires to perform operations on both phones, however, the UI does not make clear when to switch to the second phone. Consequently, several participants required help during this step because they did not know how to successfully connect the two phones. A potential solution could look as follows: as soon as the feature is activated, an illustration prompts the user to focus on the second phone. Once the phones are connected, the illustration changes and clearly informs the researcher that the connection is established and no further actions are necessary for this step. Figure 7.3 shows how such a design could guide through the process.

Lastly, participants had diverging opinions about the step for mounting the phone and aligning the camera. P5 explicitly mentioned this procedure as the most satisfying because he preferred the interaction with the physical objects over the purely on-screen settings of the previous steps.

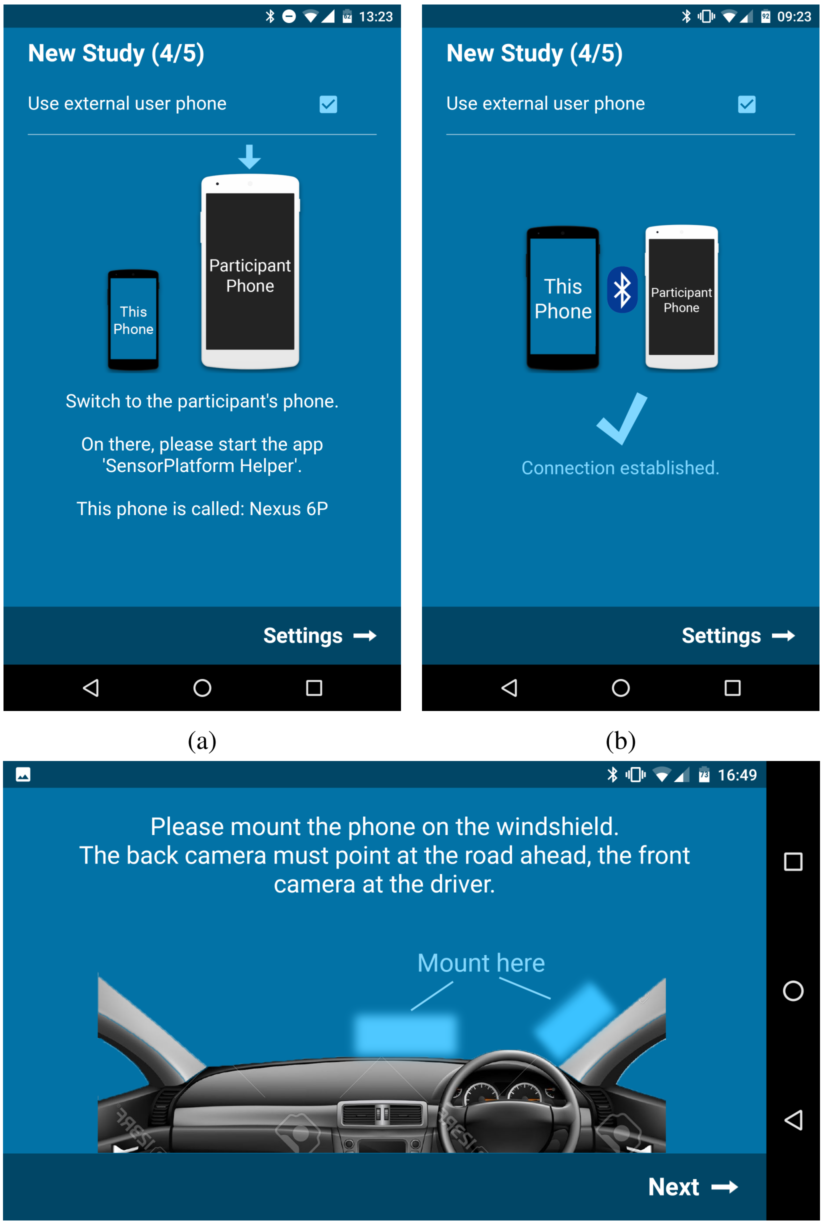


Figure 7.3: Redesigned version of the Bluetooth connection setup: **(a)** is shown at the beginning of the step; **(b)** becomes visible in case of a successful connection. **(c)** Recommendations on where to place the phone could help increasing the researcher's confidence in the setup.

On the other hand, a number of participants stated confusion about how and where to mount the phone. Again, they wished for clearer instructions or "best practices" in the form of textual hints or illustrations.

The experts' feedback revealed a critical flaw: Because the phone has to be aligned with the vehicle coordinate system in order to yield correct data, it is crucial to design this interaction in a way that results in few errors. In the evaluated version of the application, participants had difficulties finding the right spot to mount the phone and aligning the cameras while keeping the phone in a horizontal position. Similar to the suggested improvements for the previous steps, an image as shown in Figure 7.3c) pointing out the best areas to mount the phone seems to be the straightforward solution.

Coming back to underlying **research question RQ3a**, the collected feedback allows drawing a conclusion and formulating an answer:

Overall, the study's qualitative results indicate that participants were able to perform the setup procedure without major problems. The application was perceived as **simple**, **straightforward**, and – perhaps most importantly – as **time-saving** in comparison to professional DAS. At the same time, participants struggled with the more complex steps that require interactions with external hardware, most prominently connecting the second phone and installing the windshield mount. As suggested by participants, providing more detailed instructions, ideally in the form of visuals, can help to reduce the perceived complexity.

Also, it can be expected that road safety experts will increase their efficiency in the long run. All participants in this study had no prior knowledge about the application and were performing the setup for the first time. Naturally, completing the setup will take less time once they are familiar with the process. Several participants mentioned this during the study and predicted a **habituation and learning effect**. As the setup process has to be performed for each participant of a NDS, the fact that the steps are always the same should be taken into account. However, this also emphasizes that a remote setup as suggested in the focus group could speed up the process significantly.

7.4 Technical Evaluation

The question about how effectively road safety researchers can operate the application was explored in the first study of the evaluation. Equally important is an evaluation of the actual data collection abilities of the low-cost approach investigated in this thesis. According to recommendations gathered in the previous focus group (see Section 7.2), a comparison of the smartphone application to an industry-grade data acquisition system can provide the necessary information on how well the resulting data sets match. Naturally, a DAS suitable for the comparison does not necessarily have the exact same feature set as *Sensor Platform*, especially in the categories of environment data and driver data. For that reason, the focus of the comparison was limited to the vehicle data, i.e., data on position, speed, accelerations, and turns.

Comparison DAS: *VigilVanguard*
Several possible options for the comparison DAS were considered. The initial idea of leveraging a tracker for fleet management[26] was discarded because access to the raw data is only possible through the provider. In the end, the decision fell on the *VigilVanguard Driver Training System* developed by Vigil Solutions[27]. *VigilVanguard* was designed to improve driver trainings by recording raw driving data (speed, location, lateral and longitudinal accelerations, videos) and events. After a training drive, the instructor and trainee can then discuss the collected data and review specific driving situations. As such, *VigilVanguard* does not provide real-time feedback during the drive. With regard to these properties, *VigilVanguard* is a suitable DAS to evaluate the following basic data collection capabilities of *Sensor Platform*:

- Accelerometer
- GPS
- Speed
- Event Detection (Accelerations, Braking, Cornering)

VigilVanguard consists of a central sensor hub that bundles all data, a GPS sensor, an accelerometer, and four cameras. Each of these six devices is connected to the sensor hub via cables. In order to operate the system and store the data, the sensor hub must be connected to a laptop PC during the drive. All parts of *VigilVanguard* can be stored in a suitcase for safe transport (see Figure 7.4). The system used for this evaluation is not the most current version of *VigilVanguard*.

[26]GeoTab7 by http://www.securatrak.com.au/, accessed 22.11.16
[27]http://vigil-solutions.com/fire/technologies/vigil-vanguard, accessed 22.11.16

Figure 7.4: The Vigil Vanguard DAS: 1 sensor hub, 1 GPS sensor, 1 accelerometer, 4 cameras, cables for connections

7.4.1 Expertimental Setup

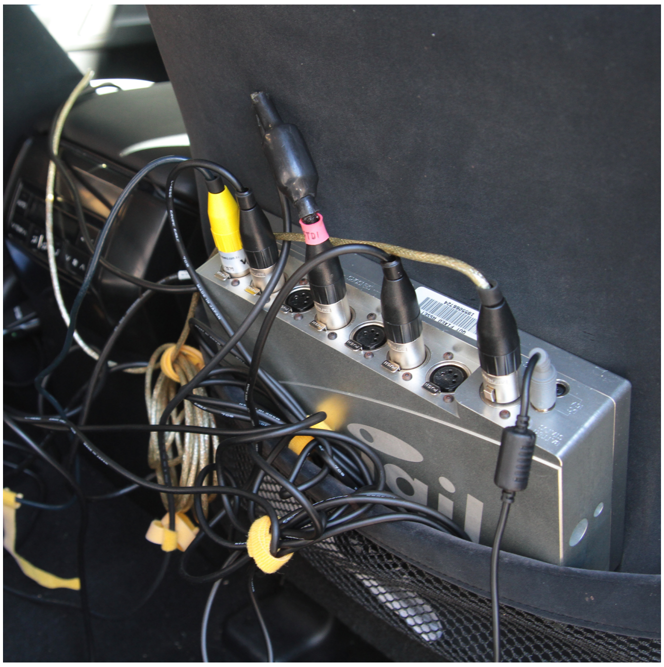
In order to compare the data collected by *Sensor Platform* and *VigilVanguard*, both systems were installed in a Toyota Prado with automatic transmission. A single test vehicle was used to limit the variance in the data. Testing across several car models would require a significantly higher sample size. Due to the preliminary nature of this study, a small sample size was deemed sufficient.

The *VigilVanguard* sensor hub was secured on the back of the driver's seat. Following the recommendations of the user manual, the GPS sensor as well as the accelerometer were mounted on the roof of the vehicle. Out of the four available cameras, only two were utilised and fixed on the windshield. One camera pointed forward to capture the road ahead whereas the second camera was focused on the driver. After connecting the sensor hub to the laptop, no additional settings were required. Figure 7.5 shows various details of the installed DAS.

In order to obtain similar video data, the smartphone running *Sensor Platform* was mounted on the windshield in a suction cup phone holder on the very right part of the vehicle's windshield. That way, the phone's back camera could capture the road while the front camera was pointed at the driver. Two external devices were connected to the application: an OBD-II adapter was plugged into the car and a Motorola Moto360 smartwatch measured the driver's heart rate. In *Sensor Platform's* study setup, the following settings were chosen:

- Sampling Rate: 4Hz
- Active Sensors: **Accelerometer**, **GPS**, **Rotation**, Cabin Light, Heart Rate, Weather, OBD-II
- Log Levels: Low, Medium, High
- Acceleration Thresholds: 0.15g, 0.25g, 0.4g
- Turn Thresholds: 0.4 rad/s, 0.55 rad/s, 0.7 rad/s
- Video Saving: **Front** and **Back** Camera
- Video Resolution: 320x240

In addition to the features needed for the direct comparison (marked bold), several other data sources such as the heart rate or speeding detection were activated in *Sensor Platform*.



(b) Cameras pointing at the road and the driver.

(a) The sensor hub.

(c) The GPS antenna on the roof.

Figure 7.5: *VigilVanguard* components: sensor hub, cameras, and GPS antenna.

Study Route

A driving route (compare Figure 7.6) was specified prior to the study drives in order to keep the driving conditions consistent. The route had a total length of 14.5km on roads through suburban areas in Brisbane, equally distributed on bigger roads with at least two lanes (7.5km) and small roads with one lane (7km). By defining a specific test route it was ensured that the same number of turns, traffic lights, and speed limits had to be passed on each drive. Start and end of each drive was a parking spot at the university's campus.

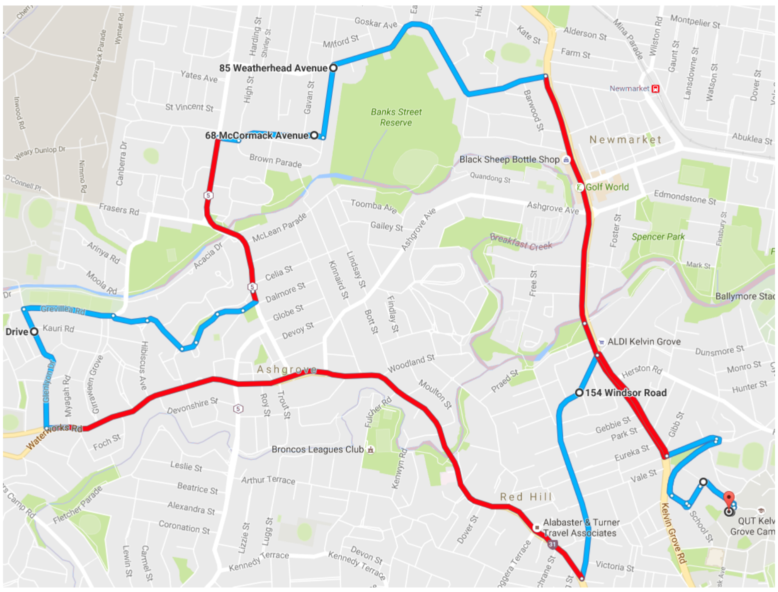


Figure 7.6: The pre-defined study route. The red parts are big roads with two or more lanes, blue routes are suburban with only one lane.

Study Procedure
Each of the experimental drives was conducted by the principal researcher, accompanied by a colleague who operated the *VigilVanguard* laptop and monitored that both systems are operational. This approach ensured that the principal researcher could entirely focus on driving. At the beginning of each drive, both researchers completed a prepared checklist in order to set up the systems. This included checking cables as well as power and data connections. The order of the steps was especially important for setting up *VigilVanguard*: for example, the system would only function correctly if the data connection was established after starting the car's engine. At the end of each drive, the battery consumption on the phone was noted down. The complete protocol can be found in Appendix E.

To counterbalance the fact that only one driver was performing the test drives, a number of trips was driven in a more **aggressive style**. Specifically, this included accelerating harshly or braking abruptly if it was safe to do so. The speed limits were followed in every drive, independent of the driving style. The purpose of variating the driving behaviour was twofold: Firstly, it can – to some extent – simulate the difference between drivers and secondly, it allows analysing how well both systems detect the increased aggressiveness in terms of event counts. With regard to the prototypal nature of the study, it can be assumed that this method is sufficient to evaluate the developed system.

7.4.2 Results

In total, the pre-defined route was driven seven times over the span of two days. On both days, the weather was dry and sunny. Due to one corrupt file on the *VigilVanguard* system, only six of the recorded drives were used for data analysis. As listed in Table 7.1, each of these drives lasted around 30 minutes, ranging from 26 to 32 minutes (M_{Time}=28.33, SD_{Time}=2.16). Drive 4 and 7 took place in the afternoon and were thereby affected by more traffic.

	Drive 1	**Drive 2**	**Drive 4**	**Drive 5**	**Drive 6**	**Drive 7**
Time in Min	29	26	32	27	27	29
Start Time	10:01	11:12	14:49	10:12	10:44	13:58
End Time	10:30	11:38	15:23	10:39	11:09	14:27
Driving Style	Normal	Normal	Normal	Aggressive	Aggressive	Normal

Table 7.1: Details about six valid study drives.

The collected raw data in the form of CSV files was taken as the basis for most of the analysis. *VigilVanguard's* sampling rate is roughly 1Hz (*Sensor Platform* 4Hz), resulting in less data rows per drive. In combination with the collected video data, the CSV files allowed taking an in-depth look at specific driving situations. Videos were not analysed in detail but were helpful in reviewing a maneuver, e.g. identifying the cause for an anomaly in the raw data. Starting with general observations, the following paragraphs present the results for GPS data, accelerometer data, turn data and event detection.

General Observations
In addition to the difference in sampling rates, several other characteristics influenced the raw data collection. On average, *Sensor Platform* collected *M*=5488 (*SD*=654.2) data rows per drive whereas *VigilVanguard* logged *M*=1683 (*SD*=136.33) rows. In three of the study drives, *Sensor Platform* detected an intermediate **end of a trip** due to exceptionally long wait times at traffic lights; the data collection resumed as soon as the vehicle started moving again. This behaviour

explains why *Sensor Platform* did log less than four times as many rows as *VigilVanguard*. On the other hand, despite the general sampling rate of 1Hz of *VigilVanguard*, a **non-linear logging behaviour** was found during the raw data analysis. In several cases, *VigilVanguard* skipped several seconds of data logging. Also, the recorded raw data seems to “jump back in time” at some points, i.e. after reporting a value for t(x), *VigilVanguard* logged the data for time t(x-5), t(x-4). The cause for this phenomenon is unknown but it affects the data quality and makes automatic analysis more difficult.

Another issue with the raw data were **inconsistencies in the OBD-II data** coming from the vehicle. Conceptually, the external OBD-II adapter was added to *Sensor Platform* in order to provide additional information sourced directly from the vehicle. The functionality was preliminary tested with an OBD-II emulator as well as a Honda Civic on the real road. However, in combination with the Toyota Prado used in this study, the OBD-II connection did not consistently deliver correct data. Periods of valid data are mixed with stretches of incorrect values. This behaviour only became apparent during the analysis as the erroneous values still fit within the overall expected range, e.g. the OBD-II adapter repeatedly reported a speed of around 40km/h while the vehicle was standing still at a traffic light. Most likely, the OBD protocol implemented by Toyota does not follow the standard. Judging by the observed behaviour, the car sometimes appends additional bytes to the sent package. Consequently, OBD-II data was not included in the analysis.

Following the requirements identified in the focus group, the study drives were also used to observe the **battery consumption** and **storage utilization** of *Sensor Platform*. Naturally, these values are highly dependent on the exact settings that are chosen for the data collection. With the settings of the study drives, *Sensor Platform* stored 12MB of CSV files and 4.71GB of video data, resulting in M_{CSV}=2MB and M_{Video}=785MB per drive. A large influence on these numbers had the log level, causing videos to be saved even for low risk events. Also, the video compression algorithm used in the application is not very efficient in terms of file size.

A detailed look at the battery consumption reveals that the low-cost application consumed $M_{Battery}$=124.3mAh during the average 28.3 minute of driving. This number is based on the value reported by the Android OS and does not include consumptions of the screen or other applications. When compared to the total battery size of the test phone (3450mAh), this accounts for a battery loss of 3.6%. If the screen is taken into account, the value rises to a relative consumption of 16.5% for each drive. The overall consumption can not be compared to *VigilVanguard* because the professional system requires a constant connection to a power source.

GPS

Across all drives, both systems captured the driven route accurately. The higher raw data sampling rate of *Sensor Platform* did not have any effect on the accuracy in this category because the phone’s GPS sensor is limited to an output of 1Hz. On a more detailed look, both systems revealed problems in the GPS data. The biggest issue were several periods where *Sensor Platform* lost the GPS signal for up to 12 seconds. Moreover, *Sensor Platform* struggled to report a constant position while the vehicle was standing still, e.g., at a traffic light. This behaviour did not occur with the *VigilVanguard* system. On the other hand, the comparison DAS did at times include a faulty offset in the data, resulting in a trajectory that did not follow the road. To illustrate these inaccuracies, Figure 7.7 plots two examples.

Directly linked to the GPS signal is the **speed calculation**. Not surprisingly, the reported speed is similar between the two systems. In Figure 7.8, a seven minute long part at the beginning of Drive 2 is shown. *Sensor Platform’s* raw data was further processed by removing outliers and applying a simple moving average whereas the *VigilVanguard* values are the direct output of the system. It can be assumed that *VigilVanguard* performs suitable postprocessing steps before logging the data.

(a) *Sensor Platform's* inaccuracy while stopped at traffic light.

(b) *Vigil Vanguard's* occasional offset in GPS trajectory.

Figure 7.7: Examples of the GPS trajectory out of the second drive. *Sensor Platform* is **blue**, *VigilVanguard* is **green**.

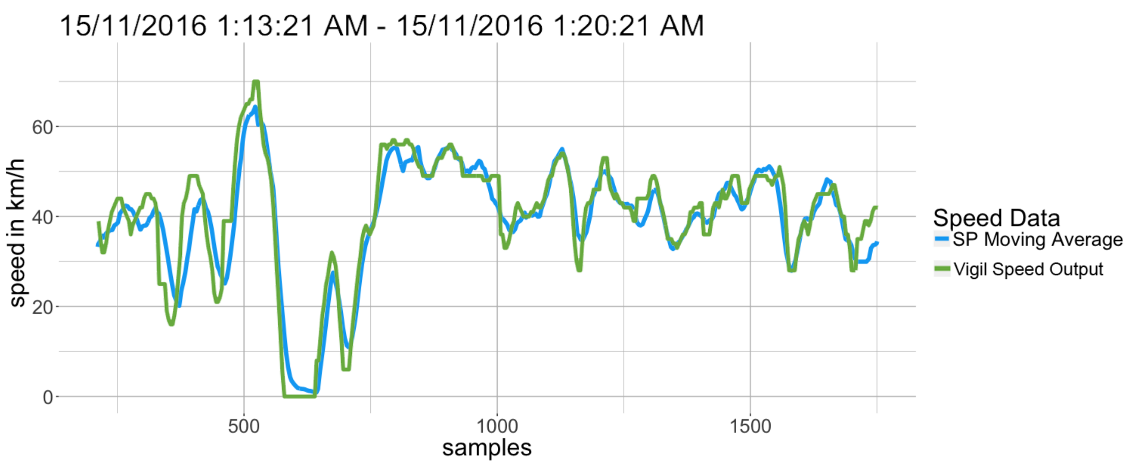


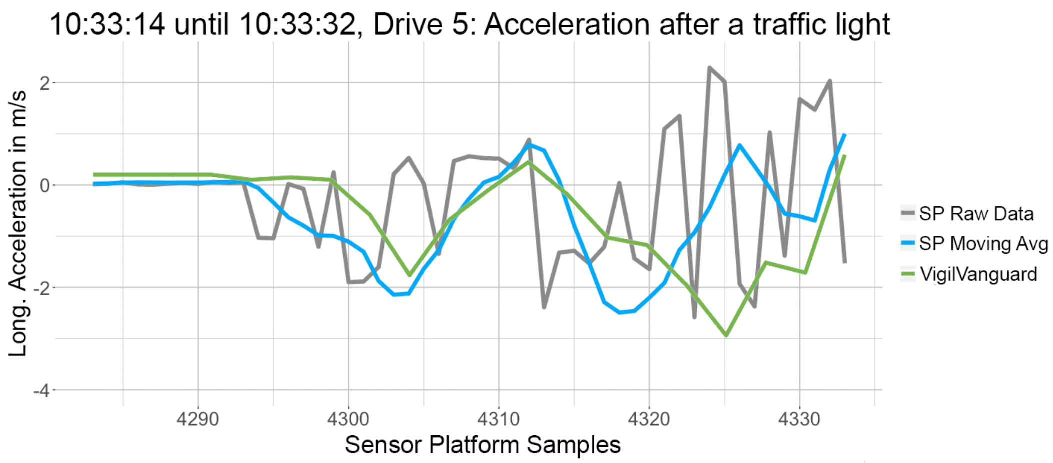
Figure 7.8: Speed data of a seven minute long part of Drive 2. Both systems produce a similar output.

Acceleration and Braking

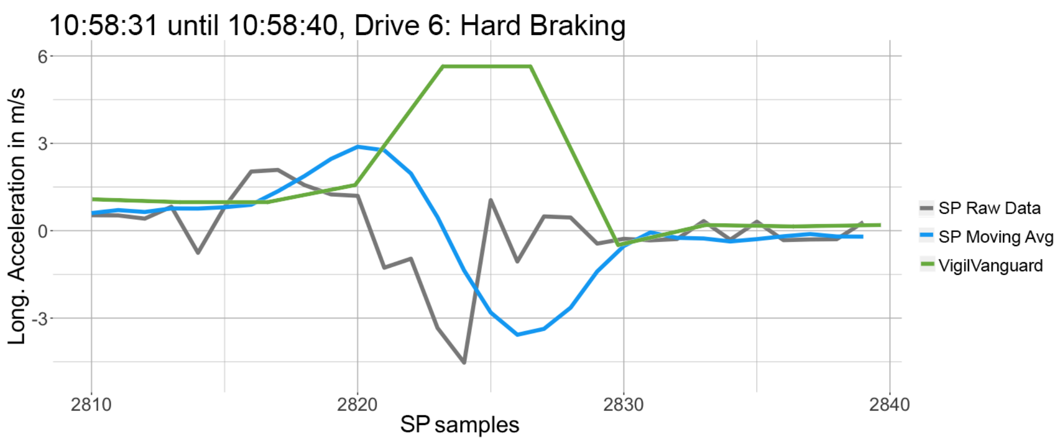
The driving style of an individual is distinguished mostly by the choice of speed and the acceleration and braking behaviour. Consequently, it is important for a data acquisition system to accurately reflect this behaviour in the collected data. *Sensor Platform* uses the phone's built-in accelerometer towards this goal whereas *VigilVanguard* can leverage a dedicated "comfort sensor" – essentially a large accelerometer – mounted on the vehicle's roof.

Overall, acceleration data collected by *VigilVanguard* more closely reflects the actual driving situations than *Sensor Platform*. The professional system outputs a very clean curve that is not visibly influenced by noisy sensor readings. In comparison, *Sensor Platform's* performance suffers from raw accelerometer data with a high amount of noise. Further processing of the data is necessary to reduce the amount of noise and thus more accurately represent the driving behaviour. An example for these observations is shown in Figure 7.9a which displays the longitudinal acceleration after a traffic light. There, the noisy raw values produced by *Sensor Platform* are clearly visible (grey curve). Yet, by applying a moving average (blue curve), the acceleration can be clearly distinguished with both systems.

Another factor influencing the accelerometer data is the vehicle's suspension. The Toyota Prado used for the study drives is suitable for off-road driving and equipped with a soft suspension. This characteristic led to notable patterns in the acceleration data of *Sensor Platform*. After an abrupt change in acceleration, e.g. a harsh brake, the soft suspension caused the vehicle to "bounce back" which triggered a peak in the opposite direction. *VigilVanguard* was not affected by this behaviour, either due to the sensor's more centered position on the roof or the more advanced filtering of the data. Figure 7.9b depicts such a harsh braking event just before a traffic light. *VigilVanguard* correctly logs the braking force whereas *Sensor Platform* is registering the particular vehicle movement caused by the suspension. This example shows that specific attributes of a vehicle can have a significant impact on the collected data.



(a) Acceleration after a traffic light. Negative values indicate an acceleration, positive values a deceleration. *Sensor Platform's* raw values (grey) need to be further filtered (blue) to get acceptable results.



(b) Accelerometer data recorded during a hard brake. *VigilVanguard* correctly registers the brake; *Sensor Platform* is influenced by the car's suspension.

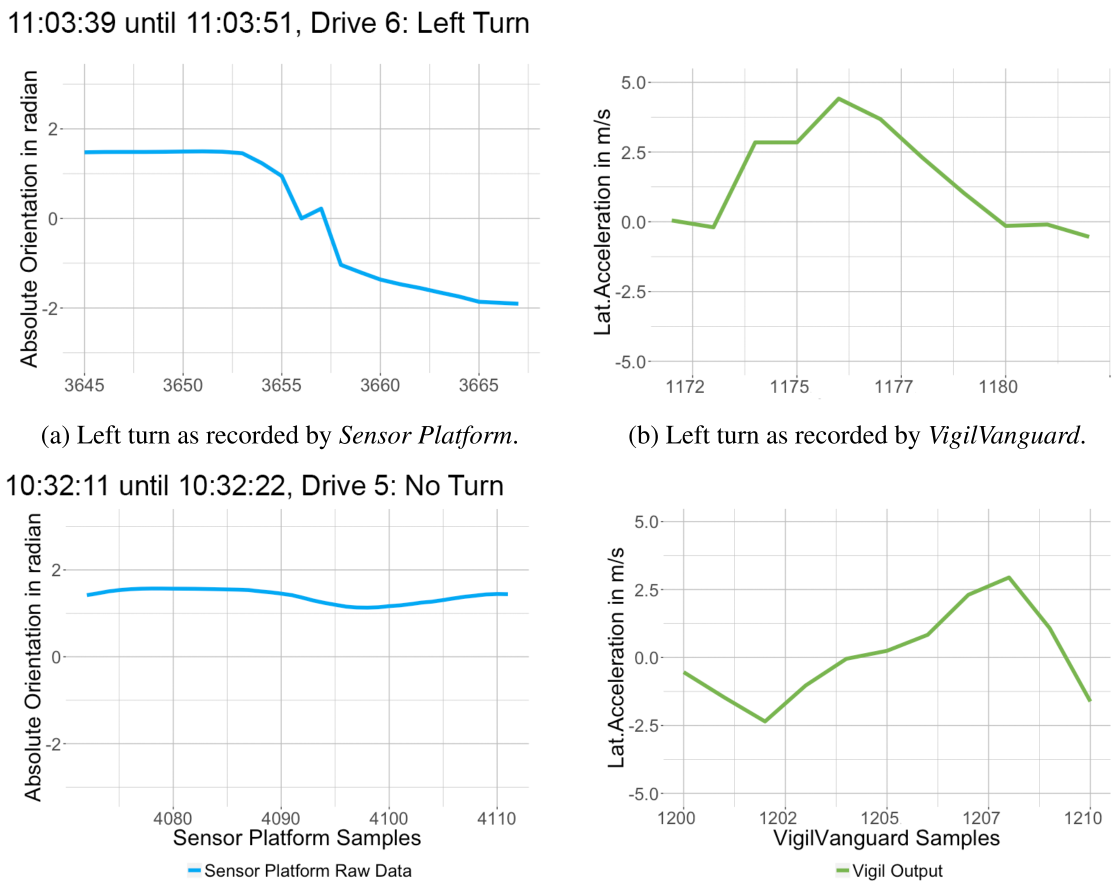
Figure 7.9: Exemplary accelerometer patterns recorded during (a) acceleration and (b) braking.

Turns

A comparison between the two systems in regard to turns is especially interesting because the underlying sensors differ. *VigilVanguard* bases the turn data on its accelerometer, the "comfort sensor", monitoring the lateral acceleration, i.e. the forces towards the left and right of the vehicle. Meanwhile, *Sensor Platform* stores the phone's absolute rotation around the y-axis, ignoring lateral accelerations.

Despite the obvious differences, both systems distinctively represent turns in the collected data. In the case of *Sensor Platform's* approach, a turn appears in the data as a transition between two constant rotation states. In contrast to that, the same turn causes a peak in the data of *VigilVanguard*, reverting back to an acceleration of zero after the turn. Figures 7.10a and 7.10b visualise these differences.

However, using the lateral acceleration to detect turns has an important disadvantage. As identified during the analysis, *VigilVanguard* often registered peaks in the lateral acceleration even in situations when the vehicle did not turn. This is caused by uneven or bumpy parts of the road (e.g. potholes). These kind of road disturbances do not result in rotations around the phone's y-axis and consequently do not appear in *Sensor Platform's* data, as evident in Figure 7.10c and 7.10d.



(a) Left turn as recorded by *Sensor Platform*.

(b) Left turn as recorded by *VigilVanguard*.

(c) Straight part on uneven road as recorded by *Sensor Platform*.

(d) Straight part on uneven road as recorded by *VigilVanguard*.

Figure 7.10: The two systems use different approaches to turn detection. *VigilVanguard* monitors lateral accelerations and is susceptible to uneven road surfaces.

Event Detection

Besides the raw data collection, both systems apply event detection algorithms to detect critical driving events. *VigilVanguard* only leverages the accelerometer and is designed to detect harsh acceleration, braking, and turns. Due to the lack of a suitable comparison, the more experimental event detection procedures of *Sensor Platform* such as tailgating or driver distraction detection were not active during the study. The following Table 7.2 compares details of the event detection systems.

	Sensor Platform	*VigilVanguard*
Event Detection Runtime	Real-time	After the drive
Detected Events	Acceleration, Braking, Turns, Speeding, Tailgating, Driver Distraction	Acceleration, Braking, Cornering
Event Levels	Low Risk, Medium Risk, High Risk	Jerky, Excessive
Thresholds	*Acceleration*: 0.15g (low), 0.25g (medium), 0.4g (high) *Turns*: 0.4rad/s (low), 0.55 rad/s (medium), 0.7 rad/s (high)	For all events: 0.3g

Table 7.2: A comparison of the two event detection systems.

For the purpose of this comparison, the events are analysed as reported by the systems. Naturally, the data sets will therefore also include false positives. However, the definition of a critical driving event is always context dependent and may vary between studies and data acquisition systems. With regard to the different thresholds and detection algorithms used, the main information of interest is not the absolute event count but whether trends are accurately reflected in the data. In the case of the presented study, such a trend that should become apparent is the variation in driving style. For this reason, the reported events were analysed without excluding possible false positives. Also, as the systems have different ways to assign levels to events (*low, medium, high* as opposed to *jerky, excessive*), the levels did not feed into the analysis.

Across all six drives, *Sensor Platform* detected an average of $M_{SP\ Event}$=60.5 (*SD*=41.58) events per drive compared to $M_{Vigil\ Event}$=84.5 (*SD*=14.07) events calculated by *VigilVanguard*. Especially noticeable is the high standard deviation of *Sensor Platform's* event count. The reason behind this finding becomes clear when the two different driving styles (normal and aggressive) are being compared (see Figure 7.11). The application developed in this thesis detected far fewer events during the normal drives, especially acceleration and braking events. Interestingly, during the last normal drive, the app registered a higher number of events which contradicts the results of the first three normal drives. As the same settings were set for each of the test drives, it is not clear what caused the application to detect this large amount of events.

In contrast to these results, *VigilVanguard* displays a more even distribution with the two aggressive drives clearly standing out. However, this distinction can only be attributed to an increase in acceleration events as the number of hard brakes and sharp turns was not influenced by varying the driving style. The high but unchanged number of detected turns from *VigilVanguard* can be explained by the false positives on uneven roads (see Section 7.4.2).

Nonetheless, in both systems the event count during the aggressive drives is significantly higher. The event count data is normally distributed as confirmed by a Shapiro-Wilk test [91] (*Sensor*

Platform: *W*=0.8778, *p*=0.2591; *VigilVanguard*: *W*=0.8949, *p*=0.3446), hence a one-way ANOVA was used to test for group differences. The analysis of variance attributes statistically significant differences between normal and aggressive driving to both *Sensor Platform* ($F(1,4)=8.79$, $p<.041$) and *VigilVanguard* ($F(1,4)=28.31$, $p<.006$). This result indicates that both the low-cost and the dedicated approach can be used to differentiate driving styles.

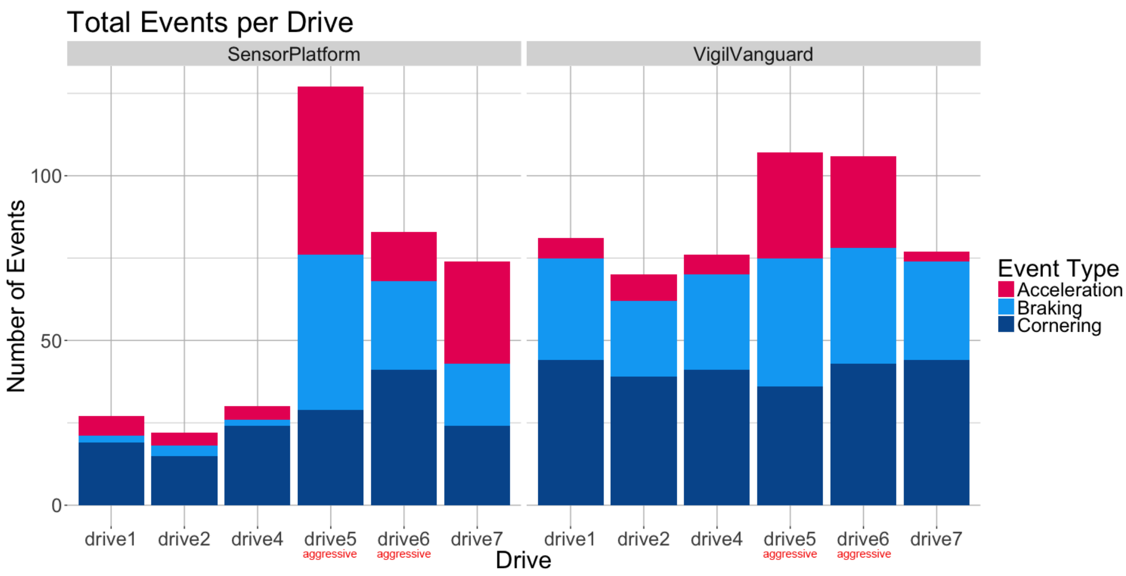


Figure 7.11: Total event counts of the two systems. Left: *Sensor Platform*, Right: *VigilVanguard*

7.4.3 Study Limitations

The technical evaluation was subject to a number of limitations, which are critically discussed in this section.

Small Sample Size: Generally speaking, the focus of the study was on identifying advantages and disadvantages of the low-cost system in comparison to the professional DAS. The analysis of the driving data focused on exemplary driving situations that highlight the issues around the different data sources. For that reason, a small sample size of N=6 was deemed sufficient. Additional drives on the same route would not provide any new insights due to the similarity of driving situations. The test drives yielded the necessary data to describe issues of the current version but do not allow generalising the data.

Single Car Model: Secondly, as the study was conducted with only a single car model, results do not include information about how the car type influences the collected data and event detection system. As pointed out in the analysis, the properties of the car such as the suspension system can have a noticeable effect. To counterbalance and further explore these effects, a greater variety of models and manufacturers would have to be tested.

Single Driver: In the same way, the driving style of each individual person is different and has an effect on the data collection. The choice to limit the amount of drivers to one was mainly due to the required ethical clearance. In Australia, any research involving the participation of humans must be undertaken in accordance with the National Statement on Ethical Conduct in Human Research[28]. A driving study on the real road requires the approval of a High Risk application that identifies any risk involved for participants. The limited timeframe of this project did not allow completing such an application. In order to address this limitation, the driver in the study varied

[28]https://www.nhmrc.gov.au/guidelines-publications/e72, Accessed 29.11.2016

the driving style between test drives (normal and aggressive). Naturally, this procedure is not able to perfectly simulate the differences in driving styles of a large set of participants but allowed estimating how the systems react to those changes.

Scope: Out of *Sensor Platform's* functionalities, only a core set was evaluated in the technical evaluation. Several features, e.g. speeding events or heart rate, were collected during the drives but not analysed because no comparison data was available. Other features such as tailgating detection and driver distraction detection were completely excluded from the technical evaluation due to the prototypal status of their implementation. Hence, the results of the evaluation do not allow drawing a conclusion on the whole feature set. Furthermore, the study was not designed to compare the systems in an absolute manner but to identify the general strengths and weaknesses of the developed prototype. After the robustness of the algorithms has been improved, future studies with a larger sample size should include the mentioned functionality and a more quantitative analysis in order to judge the complete system.

All in all, the small sample size is clearly a limitation of the existing study. Variations in the data due driving styles were simulated to a limited extent but could not be explored in detail. The influence that the car model can have on the data was acknowledged. Once the identified problem areas have been improved upon in a refined version of *Sensor Platform*, the above mentioned limitations should be addressed in a future study with a larger and more diverse sample.

7.4.4 Discussion

The comparison of the two systems was designed to yield the data that allows drawing a conclusion for **research question RQ3b**, i.e., a judgement on the accuracy of the developed smartphone DAS. As a result of the small sample size and the other limitations of the study, the systems were not compared in an absolute manner. Instead, the raw data was analysed to get a general idea about the strengths and weaknesses of the two DAS, highlighted by comparing exemplary driving situations.

In general, **driving data** collected by *Sensor Platform* is less accurate than the values recorded by an industry grade DAS such as *VigilVanguard*. The data is more easily affected by noise and environmental influences such as road conditions. Looking closer at the different categories of data, it becomes clear that most common source of error in *Sensor Platform's* data is the accelerometer. Raw readings from the built-in acceleration sensor do not always accurately reflect the recorded driving situations as the sensor is reacting to even the smallest changes caused by vibrations of the vehicle. This finding is very similar to the results of Castignani et al. [67]. Theoretically, this behaviour can be improved by the use of a low-pass filter, however, the applied algorithm was not able to filter out all unwanted signals. This indicates that *VigilVanguard* includes superior filtering algorithms that are able to reduce the impact of vibrations, resulting in a clearer representation of the actual driving behaviour.

However, not only *Sensor Platform* is affected by misinterpreted accelerometer readings. This is illustrated by the **turn detection** algorithm used in the *VigilVanguard* system: changes in road surface angles cause peaks in lateral accelerations that look similar to those recorded during actual turns. These situations were handled better by the gyroscope-based data of *Sensor Platform* as it is not affected by vehicle movements other than turns around the y-axis.

While the performance of the accelerometer differs between the two systems, other sources of data, specifically the **GPS sensor**, produce very similar results. Apart from occasional signal losses in *Sensor Platform's* data, the GPS signal filtered with a linear Kalman filter matched the driven route and speeds closely. Even more, a future version of the application could implement an extended Kalman filter which leverages data from additional sources to further increase reliability.

In light of the identified issues with the raw data, it is not surprising that the **event detection** of both systems resulted in a number of false positives and negatives. In case of *Sensor Platform*, most of these were due to misinterpreted longitudinal accelerometer data whereas *VigilVanguard* is affected by falsely detected cornering events. The consequences of these inaccuracies are dependent on the specific use case: If the aim is to detect general trends in driving styles, both systems' data clearly reflect the differences between normal and aggressive driving. However, if an exact event count is needed, a trained operator is required to review all collected events to filter out false positives. Looking at previous large scale NDS, the manual review of events is part of the procedure: "Video is also necessary to allow for screening the invalid events that will inevitably be collected." [4, p.329].

In contrast to the obvious limitations in terms of data accuracy, a major advantage of *Sensor Platform* is the **greater variety of data sources**. In addition to the compared sensors, researchers can analyse data on speed limits, heart rate, and internal car values such as engine speed and fuel consumption. However, the technical evaluation revealed a number of issues with these sources that need to be addressed in future developments: the smartwatch failed to deliver a continuous output of the heart rate as it is susceptible to losing contact to the skin. Even more, the physiological data from consumer devices is of questionable reliability (see [35]). Also, the malfunction of the OBD-II adapter during the study highlights the need for an option to customise command codes.

7.5 Recommendations

The two parts of the evaluation as described in Sections 7.3 and 7.4 examined the developed application from an user and a technical perspective. On basis of the findings, a number of recommendations stand out for improving *Sensor Platform*. Even though these guidelines were derived from the evaluation of a single prototype, they were generalised to apply to any mobile data collection system.

- **Reduce importance of the accelerometer.** The technical evaluation made evident that the built-in accelerometer is not a perfect fit for an accurate representation of driving data. While overall trends are visible, the data is too susceptible to other environmental influences and reacts to every small vibration of the vehicle. More advanced filtering techniques (e.g. Kernel Adaptive Filtering [92]) have the potential to remove the environmental noise to some extent. However, it can be assumed that due to the lower quality of the phone's built-in sensor, the accelerometer will always produce less accurate results than a dedicated sensor as used in *VigilVanguard*. For that reason, the accelerometer should not be used as a central data source. Instead, other sensors can be leveraged to cover the same functionality.

- **Put more emphasis on the vehicle speed.** The vehicle speed stands out as an option to replace the accelerometer data. Since acceleration is a change in velocity over time, it can be derived from the previous speed readings. Of course, the current GPS-based method to calculate speed at a rate of 1Hz is not detailed enough for accurate acceleration estimates. However, methods like extended Kalman Filtering allow extrapolating positions to simulate a higher sampling rate. Alternatively, the on-board diagnostics adapter can be used to acquire accurate speed values from the vehicle. Other research projects and applications came to a similar conclusion: Castignani et al. [67] report noisy accelerometer values and introduce an approach to fuse GPS speed and acceleration to counterbalance the noise effects. Another example is *DriveFlo*[29], a phone application that scores driving behaviour, which relies solely on the GPS speed to detect events such as hard accelerations or brakes. These examples show that the vehicle speed can be used to calculate accelerations and detect events as sudden changes in speed.

- **Tap the potential of the OBD interface.** In the prototype version developed in this thesis, the external OBD-II adapter was used to gather data on speed, engine speed, and fuel con-

[29] https://www.driveflo.com/flo-driving-behavior-analysis-how-does-it-work/, Accessed 28.11.2016

sumption. Despite the malfunction during the technical evaluation, this connection has a lot of untapped potential. For example, information about the steering wheel angle can replace the gyroscope based turn detection. Even more, pedal use data can be fused with the vehicle speed to detect accelerations and brakes, unaffected by accelerometer noise. The difficulty in accessing these values is the heterogeneity of car models. Most manufacturers loosely follow the OBD standard but also include their own command codes. The issue becomes apparent in the malfunction during the study: command codes that previously worked with one car model failed to consistently query valid data in another vehicle. On account of this problem, a future version of *Sensor Platform* should include the option to customise command codes for specific car models. That way, researchers can choose which values they want to access and can provide the correct codes.

- **Avoid purely textual instructions during the study setup.** The main finding of the expert evaluation is that users tend to not read instructions when they are presented as text. This behaviour caused experts during the study get stuck in the setup process, not knowing what to do next. Even more critical are errors in mounting the phone as they may render the collected data unusable. In order to avoid such situations, the application should visualise the required actions in addition to the textual instruction. Furthermore, changes in the application's state should be highlighted through visualisations, too. In Section 7.3.3, Figure 7.3 exemplifies how this recommendation could be implemented.

8 Conclusion

This chapter takes a look back at the project and assesses the obtained findings by answering the last research question. It proceeds with a summary of the thesis, followed by a short discussion of the project's contributions. Finally, the chapter concludes with an outlook on how this concept could further develop in the long term, proposing extensions and avenues for future research.

The answer to **research question RQ4** is sourced from the combined findings of all the previous chapters. Throughout the different phases of this thesis, the drawbacks and benefits of the low-cost approach became apparent.

Disadvantages
Most notably among the disadvantages is the **limited accuracy** of the collected data. In the current implementation, the app's data collection was affected by sensor noise and environmental influences. Although improved filtering algorithms can counterbalance this issue to some extent, high-cost professional systems will always have the advantage of using dedicated hardware with software tailored to it.

With regard to **functional range**, smartphone-based systems lack some functionality that is usually present in professional setups, especially data produced by a radar. These sensors provide reliable information about other road users in the proximity – an essential part of environment data. Previous NDS often included several radars (compare Table 3.1). On the smartphone, this type of data can currently only be extracted from the camera feed, which is much more limited in field of view and range. In the future, affordable and lightweight radars such as the *Garmin Varia*[30] might open new possibilites to low-cost systems.

Benefits
On the other hand, the developed smartphone-based DAS revealed a number of benefits. As smartphones are ubiquitous devices, the majority of road safety researchers is already familiar with how to use them. If this advantage is combined with a **good usability** of the data collection app, the approach results in an easy-to-use system, as became evident in the expert user study in Chapter 7.3. In comparison, the tested professional system requires a much more complex installation in the car and additional training to operate the software.

While some features of traditional DAS are missing, smartphones can introduce a range of **new functionality** to NDS that was previously hard to include. Most notably is the ability to monitor the driver's personal phone. As distraction caused by mobile phones was often mentioned by experts in the interviews (see Chapter 5.1) and is a frequent topic in literature (e.g. [73, 93]), this feature offers valuable insights into driver behaviour. Furthermore, physiological devices such as smartwatches add another dimension to the category of driver data. Finally, the open-source software allows extending the functional range. This is in most cases not possible when using a traditional off-the-shelf DAS.

Obviously, the **involved costs** are another important advantage. The mobile data collection system promises to reduce the costs from several thousand dollars to the price point of well below 1000$ per participant. For example, the hardware used for *Sensor Platform* can be bought for about 750$ on the date of writing. In addition to that, the smartphone system is easy to install in a car, thus no trained technician is needed for the setup. At that cost, naturalistic driving studies can be arranged with a much larger number of participants. Even though some data accuracy will be sacrificed by using a smartphone application, the large sample size can ideally counterbalance this drawback.

In the end, the question of which approach is better suited is dependent on the requirements of a specific research project. If the highest possible accuracy is required, the smartphone app is not

[30]https://buy.garmin.com/en-US/US/p/518151, Accessed 13.02.2017)

yet able to compete with industry-grade systems such as the one from VTTI[31]. However, if the study is of an explorative nature and aims to detect trends rather than absolute values, the low-cost approach is well worth considering, as the above mentioned advantages show.

8.1 Summary

Naturalistic driving studies have become a popular method in road safety. These studies yield unrivaled insights into how people actually behave during everyday drives. However, the equipment necessary to conduct NDS comes at a very high cost; this fact makes real-road evaluations expensive and poses a large overhead in administration.

In my thesis, I aimed to develop a more affordable way to conduct naturalistic driving studies. In particular, I explored the design of a smartphone application. Through methods of user-centered design, the requirements for such an application were defined in close cooperation with road safety experts.

The resulting prototype, the Android application *Sensor Platform*, leverages the phone sensors as well as external devices in order to collect vehicle, driver, and environment data. In an evaluation consisting of two parts, I tested the application from a user's perspective as well as from a technical point of view. While the ten expert users liked the general usability and provided positive feedback, test drives revealed a number of accuracy issues in the collected data. In contrast to that, *Sensor Platform* is able to deliver functionality that goes beyond traditional DAS, i.e. physiological measures and logs of phone interactions. With these findings in mind, I further derived several recommendations that promise to improve the accuracy and robustness of data collection. Future versions of *Sensor Platform* as well as related projects can build on these guidelines.

8.2 Contributions

By exploring the four research questions, this thesis contributes to future research in several ways. Besides the recommendations given in Chapter 7.5, the two main contributions are as follows:

Most importantly, the project explored the design of a mobile DAS by explicitly targeting **road safety experts as a user group**. This contrasts existing mobile data collection tools that often aim to provide feedback to the individual driver. Through methods of user-centered design, the identified requirements fit the specific needs of experts and are tailored to support studies on the real road. This focus is highlighted by features such as the post-drive questionnaire or *Sensor Platform's* ability to log phone interactions.

This feature of monitoring phone interactions while driving is a novelty among data acquisition systems and sets *Sensor Platform* apart. As driver distraction is one of the most common causes for accidents [4], road safety experts are very interested in understanding how and when drivers interact with mobile devices. In contrast to camera based evidence, the log files of *Sensor Platform* allow analysing not only the duration but also the type of interaction. While efforts have been made to include phone log files in previous NDS (e.g. [73]), the collected data did only take calls and text messages into account. Of course, these represent only a small percentage of phone interactions that happen during a drive. Potentially more dangerous are social media apps or games. From a road safety perspective, it is a major difference whether the driver is interacting with a navigational software or writing messages in a chat app. *Sensor Platform* allows logging screen touches in combination with the name of the visible application, therefore enabling researchers to look at different categories of applications and their effects on driving style.

The second contribution is made in the form of the **publicly available data collection service** that runs on Android. It supports faster prototyping as future applications can simply build on top of

[31] http://www.vtti.vt.edu/facilities/labs-das.html, Accessed 19.01.2016

the service and process the collected driving data. Benefits are expected not only in context of naturalistic driving research but also in different areas of road safety research. For example, the data collection service could be used to evaluate concepts for safety interventions on the real road. In the future, applications such as *CoastMaster* [9] which have previously only been tested in driving simulators can take the next step and deploy the concept in natural driving environments. Moreover, the service can be extended to support new types of sensors, e.g. advanced physiological measuring devices. In conclusion, the freely available data collection service holds the potential to support future projects in investigating a broad range of research questions.

8.3 Outlook

The developed low-cost data acquisition system, *Sensor Platform*, has two main areas that can be improved upon in future work: the **functional range** and the **algorithms** used to process the real-time data.

Even though I defined *Sensor Platform's* **functional range** in close cooperation with road safety experts, not all of the discussed features could be included in the prototype. Apart from obvious features such as lane detection, the participants of the focus group named a number of extensions that they are interested in. Some of these features are not part of the Android application but can improve the remote administration of a driving study.

Clearly, the development of a web server to store and analyse the data was not the main focus of this thesis. However, it is an important part of a complete data acquisition system and presents an interesting area of research for future work. Projects could explore questions from data analysis and information visualisation: How can we automatically condense the large amount of driving data? What kind of visualisations can support road safety experts in gathering insights?

The second area of improvement are the **algorithms** that collect and process the raw data on the phone. In the current implementation, the focus was on providing a toolbox and not on finding the perfect filtering methods for one type of data. As became evident in the technical evaluation, this approach naturally results in some inaccuracies of the collected data. Taking *Sensor Platform* as a basis to build upon, future projects can now focus on specific algorithms to increase the accuracy.

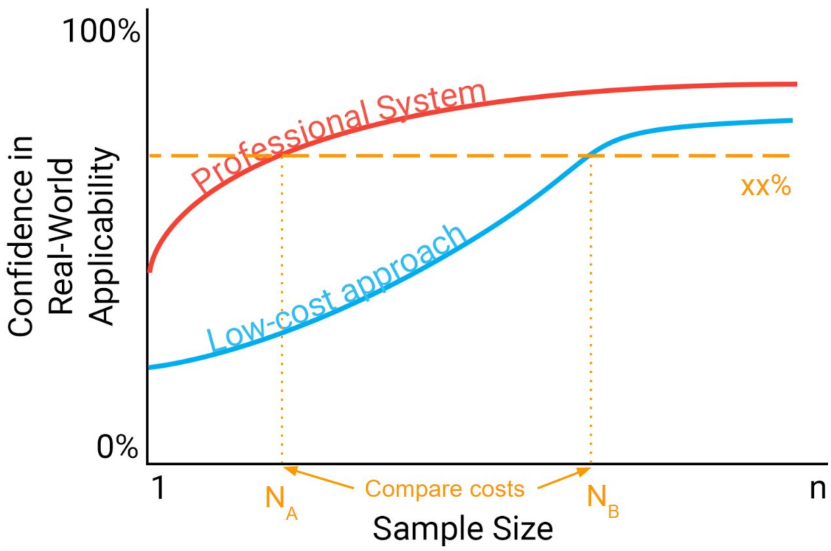


Figure 8.1: An approach to comparing two systems: to reach a confidence of x%, the professional system needs a sample size of N_A, whereas the low-cost systems requires N_B participants. For these two sample sizes, the costs can be compared to inform a decision.

Nevertheless, a low-cost DAS on basis of a smartphone will likely never have the same level of accuracy as a professional system with dedicated hardware. Road safety researchers who want to weigh both options could compare two systems as depicted in Figure 8.1. The professional system might reach the desired “confidence level”, i.e. how well the data represents reality, with a comparatively small number of participants whereas the low-cost DAS requires a large sample size to reach the same level of confidence. Then, the actual costs for buying and installing the required number of units can be compared. Of course, this method is simplified and needs a way to reliably assess the confidence level of a data set.

Despite the known drawbacks, mobile data collection systems will play a major role in road safety research in the coming years. Low costs paired with the opportunity to reach a large number of participants are essential advantages. As more time and expertise is invested in the development of these mobile systems, the accuracy and functional range will further increase, decreasing the gap to more expensive hardware. The insights gathered in naturalistic studies will ultimately help to increase safety on the roads. Possibly, future safety interventions can then be tailored to each driver on the basis of previously collected data, thus adapting to individual driving styles.

A Road Safety Expert Interview

Demographics

Interview Number:
Age: year
Gender: f / m

First, I would like to talk briefly about your area of research. Could you tell me what you are currently working on?

1. What aspects of driving/driver behaviour are most important to your research? (e.g. driver inattention, aggression, mobile phone use)

   And in general?

2. What kind of driving data are you interested in?

3. Not thinking about feasibility or limitations, what data would be great to have?

4. What are relevant driving / driver states?

5. What are "critical driving events" that need to be detected?

6. Are there any specific devices or sensors that you would like to use? (sensor not installed in a smartphone, e.g. physiological signals, temperature sensor)

7. Do you know any apps designed to be used in the car?

   If yes, did you use them?

   If you did use one of them, what did you like / did not like about it?

   If no, why did you decide not to use them? What would have to be different for you to use them?

8. Can you think of any other requirements the system should fulfill? (technical or non-technical)

9. What are current trends in road safety research? How will it change in the future? What will be the main challenges?

10. Additional notes

B Expert Evaluation App Summary

Sensor Platform is an Android application that can be used as a data acquisition system in real-world driving studies. Simply set up a new study, select the values you are interested in and start data collection! All collected data will be stored in log files for later analysis. Additionally, short video clips are recorded around critical driving events.

SENSORS
The app uses smartphone sensors and cameras to record various driving data values, e.g.

- Acceleration
- Location
- Speed
- Weather
- Cabin Light
- Following Distance

EVENTS
Furthermore, the app analyses the data in real-time to detect critical driving events such as

- Hard accelerations and braking
- Cornering
- Speeding
- Driver Distraction
- Tailgating
- Phone interactions

PLUGINS
In addition to the internal sensors, the app can connect to external sensors via Bluetooth to gather even more data.

- OBD-II Adapter: collect accurate speed, fuel consumption and engine load
- User Smartphone: connect the user's personal phone and monitor if and how the driver interacts with it while driving
- Smartwatch: log the driver' heart rate

C Expert Evaluation Tasks

TASK 1:
Start the application *Sensor Platform* on the research phone.
Create a new study and add a new participant:
Study Name: “myNDS”
Study ID: “001”
Participant ID: “p007”
23 years, male

When you are done, press “Sensors” at the bottom to continue.

TASK 2:
Activate the following sensors:

- Accelerometer
- GPS
- Rotation
- Front Camera
- Back Camera
- Speeding Detection

When you are done, press “OBD-II” at the bottom to continue.

TASK 3:
Plug the OBD-II adapter into the car and connect the application to it.

When you are done, press “User Phone” at the bottom to continue.

TASK 4:
Connect the application to the participant’s personal phone.

When you are done, press “Settings” at the bottom to continue.

TASK 5:
Take some time to explore the different settings on this screen.
What do you discover?
Let me know when you are finished.

Followed by question: “Which settings influence the amount of data being stored?”

TASK 6:
Follow the instructions to mount the phone and start the study.

D System Usability Scale

SYSTEM USABILITY SCALE

Participant ________
Age ______________
Date ______________

	Strongly disagree				Strongly agree
1. I think that I would like to use this system frequently	☐	☐	☐	☐	☐
2. I found the system unnecessarily complex	☐	☐	☐	☐	☐
3. I thought the system was easy to use	☐	☐	☐	☐	☐
4. I think that I would need the support of a technical person to be able to use this system	☐	☐	☐	☐	☐
5. I found the various functions in this system were well integrated	☐	☐	☐	☐	☐
6. I thought there was too much inconsistency in this system	☐	☐	☐	☐	☐
7. I would imagine that most people would learn to use this system very quickly	☐	☐	☐	☐	☐
8. I found the system very cumbersome to use	☐	☐	☐	☐	☐
9. I felt very confident using the system	☐	☐	☐	☐	☐
10. I needed to learn a lot of things before I could get going with this system	☐	☐	☐	☐	☐

Sensor Platform

E Technical Evaluation Checklist

1. Power-plug in the trunk is plugged in
2. All sensors are connected to the sensor hub
3. Vigil cameras suction cups are securely mounted
4. Sensor hub has power -> green light
5. Laptop is connected to power
6. Laptop is connected to Sensor hub and detects cameras
7. Smartwatch is on wrist
8. Sensor Platform is set up and mounted: "Waiting for trip to start"
9. Start car engine
10. Start VigilVantage Software
11. Note down smartphone battery percentage
12. VigilVantage "Start recording"

——————- DRIVING ——————-

12. Before turning engine off: VigilVantage "Stop recording"
13. Sensor Platform "Stop"
14. Engine off
15. Note down smartphone battery percentage

List of Figures

LIST OF FIGURES

Content of the attached CD

/01_Document ....The thesis in PDF and LaTex format

/02_References ....All available references as a digital copy

/03_SourceCode ....The source code of *Sensor Platform*
   /01_SensorPlatform
   /02_Server

/04_StudyData ....Raw data files of all studies
   /01_Interviews
   /02_FocusGroup
   /03_UserStudy
   /04_Technical

/05_Presentations ....Presentations about the thesis project

References

[1] Department of Violence and Injury Prevention and Disability, “Global status report on road safety 2015,” World Health Organization (WHO), Tech. Rep., 2015.

[2] D. Mohan, “Road safety in less-motorized environments: future concerns,” *Int. J. Epidemiol.*, vol. 31, no. 3, pp. 527–532, Jun. 2002.

[3] J. R. Treat, N. S. Tumbas, S. T. McDonald, D. Shinar, R. D. Hume, R. E. Mayer, R. L. Stansifer, and N. J. Castellan, “Tri-Level study of the causes of traffic accident. volume i: Causal factor tabulations and assessments,” U.S. Department of Transportation, Tech. Rep., 1977.

[4] T. A. Dingus, S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. D. Sudweeks, M. A. Perez, J. Hankey, D. J. Ramsey, S. Gupta, and Others, “The 100-car naturalistic driving study, phase II-results of the 100-car field experiment,” Tech. Rep., 2006.

[5] D. A. Johnson and M. M. Trivedi, “Driving style recognition using a smartphone as a sensor platform,” in *14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2011.

[6] C.-W. You, N. D. Lane, F. Chen, R. Wang, Z. Chen, T. J. Bao, M. Montes-de Oca, Y. Cheng, M. Lin, L. Torresani, and A. T. Campbell, “CarSafe app: Alerting drowsy and distracted drivers using dual cameras on smartphones,” in *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys ’13. New York, NY, USA: ACM, 2013, pp. 13–26.

[7] J.-H. Hong, B. Margines, and A. K. Dey, “A smartphone-based sensing platform to model aggressive driving behaviors,” in *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI ’14. New York, NY, USA: ACM, 2014, pp. 4047–4056.

[8] R. Araújo, Â. Igreja, R. de Castro, and R. E. Araújo, “Driving coach: A smartphone application to evaluate driving efficient patterns,” in *Intelligent Vehicles Symposium (IV), 2012 IEEE*. ieeexplore.ieee.org, Jun. 2012, pp. 1005–1010.

[9] F. Steinberger, P. Proppe, R. Schroeter, and F. Alt, “CoastMaster: An ambient speedometer to gamify safe driving.” eprints.qut.edu.au, 24 Oct. 2016.

[10] J. Preece, H. Sharp, and Y. Rogers, *Interaction Design: Beyond Human-Computer Interaction*. Wiley, 2007, vol. 4th Edition.

[11] “sensor,” http://www.merriam-webster.com/dictionary/sensor, 2015, accessed: 2016-8-22.

[12] D. L. Hall, *Mathematical Techniques in Multisensor Data Fusion*. Norwood, MA, USA: Artech House, Inc., 1992.

[13] G. Welch and G. Bishop, “An introduction to the kalman filter,” Department of Computer Science, University of North Carolina, Tech. Rep., 2000.

[14] Leonard A. McGee and Stanley F. Schmidt, “Discovery of the kalman filter as a practical tool for aerospace and industry,” NASA, Tech. Rep., Nov. 1985.

[15] A. D’Amico and C. D. Natale, “A contribution on some basic definitions of sensors properties,” *IEEE Sens. J.*, vol. 1, no. 3, pp. 183–190, Oct. 2001.

[16] H. J. Lee and S. Jung, “Gyro sensor drift compensation by kalman filter to control a mobile inverted pendulum robot system,” *Industrial Technology, 2009. ICIT 2009. IEEE*, 2009.

[17] P. A. Zandbergen, “Accuracy of iphone locations: A comparison of assisted GPS, WiFi and cellular positioning,” *Trans. GIS*, vol. 13, pp. 5–25, 1 Jun. 2009.

[18] E. Farmer and E. G. Chambers, “A study of personal qualities in accident proneness and proficiency,” *Industrial Health Research Board*, 1929.

[19] M. Ishibashi, M. Okuwa, S. Doi, and M. Akamatsu, “Indices for characterizing driving style and their relevance to car following behavior,” in *SICE, 2007 Annual Conference*, 2007, pp. 1132–1137.

[20] D. J. French, R. J. West, J. Elander, and J. M. Wilding, “Decision-making style, driving style, and self-reported involvement in road traffic accidents,” *Ergonomics*, vol. 36, no. 6, pp. 627–644, Jun. 1993.

[21] J. Reason, A. Manstead, S. Stradling, J. Baxter, and K. Campbell, “Errors and violations on the roads: a real distinction?” *Ergonomics*, vol. 33, no. 10-11, pp. 1315–1332, Oct. 1990.

[22] D. Parker, S. G. Stradling, and A. S. R. Manstead, “Modifying beliefs and attitudes to exceeding the speed limit: An intervention study based on the theory of planned behavior,” *J. Appl. Soc. Psychol.*, vol. 26, no. 1, pp. 1–19, 1 Jan. 1996.

[23] J. C. F. de Winter and D. Dodou, “The driver behaviour questionnaire as a predictor of accidents: a meta-analysis,” *J. Safety Res.*, vol. 41, no. 6, pp. 463–470, Dec. 2010.

[24] A. af Wåhlberg, L. Dorn, and T. Kline, “The manchester driver behaviour questionnaire as a predictor of road traffic accidents,” *Theoretical Issues in Ergonomics Science*, vol. 12, no. 1, pp. 66–86, 2011.

[25] R. West, D. French, R. Kemp, and J. Elander, “Direct observation of driving, self reports of driver behaviour, and accident involvement,” *Ergonomics*, vol. 36, no. 5, pp. 557–567, May 1993.

[26] D. L. Thomas and E. Diener, “Memory accuracy in the recall of emotions,” *J. Pers. Soc. Psychol.*, vol. 59, no. 2, p. 291, Aug. 1990.

[27] N. Kaptein, J. Theeuwes, and R. V. D. Horst, “Driving simulator validity: Some considerations,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1550, pp. 30–36, 1996.

[28] D. Hallvig, A. Anund, C. Fors, G. Kecklund, J. G. Karlsson, M. Wahde, and T. Åkerstedt, “Sleepy driving on the real road and in the simulator—a comparison,” *Accid. Anal. Prev.*, vol. 50, pp. 44–50, 2013.

[29] B. Reimer and B. Mehler, “The impact of cognitive workload on physiological arousal in young adult drivers: a field study and simulation validation,” *Ergonomics*, vol. 54, no. 10, pp. 932–942, Oct. 2011.

[30] J. A. Healey and R. W. Picard, “Detecting stress during real-world driving tasks using physiological sensors,” *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 2, pp. 156–166, Jun. 2005.

[31] P. M. Valero-Mora, A. Tontsch, R. Welsh, A. Morris, S. Reed, K. Touliou, and D. Margaritis, “Is naturalistic driving research possible with highly instrumented cars? lessons learnt in three research centres,” *Accid. Anal. Prev.*, vol. 58, pp. 187–194, 2013.

[32] E. K. Choe, N. B. Lee, B. Lee, W. Pratt, and J. A. Kientz, “Understanding quantified-selfers’ practices in collecting and exploring personal data,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI ’14. New York, NY, USA: ACM, 2014, pp. 1143–1152.

[33] M. Swan, “Emerging patient-driven health care models: an examination of health social networks, consumer personalized medicine and quantified self-tracking,” *Int. J. Environ. Res. Public Health*, vol. 6, no. 2, pp. 492–525, Feb. 2009.

[34] J. Kopp, “Self-monitoring: A literature review of research and practice,” *Soc. Work Res. Abstr.*, vol. 24, no. 4, pp. 8–20, 21 Dec. 1988.

[35] F. Steinberger, R. Schroeter, and D. Babiac, “Engaged drivers, safe drivers – leveraging Real-Time driving data gathered from mobile and wearable devices,” in *Automotive User Interfaces - Creating Interactive Experiences in the Car*, G. Meixner and C. Mueller, Eds. Springer, 2017.

[36] M. Swan, “Sensor mania! the internet of things, wearable computing, objective metrics, and the quantified self 2.0,” *Journal of Sensor and Actuator Networks*, vol. 1, no. 3, pp. 217–253, 8 Nov. 2012.

[37] I. Li, A. Dey, and J. Forlizzi, “A stage-based model of personal informatics systems,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI ’10. New York, NY, USA: ACM, 2010, pp. 557–566.

[38] K. Kunze, M. Iwamura, K. Kise, S. Uchida, and S. Omachi, “Activity recognition for the mind: Toward a cognitive“ quantified self”,” *Computer*, 2013.

[39] S. Hachem, G. Mathioudakis, A. Pathak, V. Issarny, and others, “Sense2health: A quantified self application for monitoring personal exposure to environmental pollution,” *SENSORNETS*, 2015.

[40] P. B. Shull, W. Jirattigalachote, M. A. Hunt, M. R. Cutkosky, and S. L. Delp, “Quantified self and human movement: a review on the clinical impact of wearable sensing and feedback for gait analysis and intervention,” *Gait Posture*, vol. 40, no. 1, pp. 11–19, 6 Apr. 2014.

[41] A. Shalom Hakkert and V. Gitelman, “Thinking about the history of road safety research: Past achievements and future challenges,” *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 25, Part B, pp. 137–149, 2014.

[42] L. M. Bergasa, D. Almería, J. Almazán, J. J. Yebes, and R. Arroyo, “DriveSafe: An app for alerting inattentive drivers and scoring driving behaviors,” in *2014 IEEE Intelligent Vehicles Symposium Proceedings*. ieeexplore.ieee.org, Jun. 2014, pp. 240–245.

[43] L. Jiang, X. Chen, and W. He, “SafeCam: Analyzing intersection-related driver behaviors using multi-sensor smartphones,” in *2016 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. ieeexplore.ieee.org, Mar. 2016, pp. 1–9.

[44] H. Han, J. Yu, H. Zhu, Y. Chen, J. Yang, Y. Zhu, G. Xue, and M. Li, “SenSpeed: Sensing driving conditions to estimate vehicle speed in urban environments,” in *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*. ieeexplore.ieee.org, Apr. 2014, pp. 727–735.

[45] A. Sathyanarayana, S. Amardeep, S. O. Sadjadi, and J. H. L. Hansen, "Leveraging sensor information from portable devices towards automatic driving maneuver recognition," in *2012 15th International IEEE Conference on Intelligent Transportation Systems*, 2012.

[46] R. Meng, C. Mao, and R. R. Choudhury, "Driving analytics: Will it be OBDs or smartphones?" 2015.

[47] J. Dai, J. Teng, X. Bai, Z. Shen, and D. Xuan, "Mobile phone based drunk driving detection," in *2010 4th International Conference on Pervasive Computing Technologies for Healthcare*. ieeexplore.ieee.org, Mar. 2010, pp. 1–8.

[48] T. Victor, J. Bärgman, M. Hjälmdahl, K. Kircher, E. Svanberg, S. Hurtig, H. Gellerman, and F. Moeschlin, "Sweden-michigan naturalistic field operational test (semifot) phase 1: Final report," *SAFER Report*, vol. 2, 2010.

[49] H. Eren, S. Makinist, E. Akin, and A. Yilmaz, "Estimating driving behavior by a smartphone," in *Intelligent Vehicles Symposium (IV), 2012 IEEE*. ieeexplore.ieee.org, Jun. 2012, pp. 234–239.

[50] M. Van Ly, M. Sujitha, and M. M. Trivedi, "Driver classification and driving style recognition using inertial sensors," in *2013 IEEE Intelligent Vehicles Symposium (IV)*, 2013.

[51] C.-W. You, M. Montes-de Oca, T. J. Bao, N. D. Lane, H. Lu, G. Cardone, L. Torresani, and A. T. Campbell, "CarSafe: A driver safety app that detects dangerous driving behavior using dual-cameras on smartphones," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, ser. UbiComp '12. New York, NY, USA: ACM, 2012, pp. 671–672.

[52] O. Kumtepe, G. B. Akar, and E. Yuncu, "Driver aggressiveness detection via multisensory data fusion," *J. VLSI Signal Process. Syst. Signal Image Video Technol.*, vol. 2016, no. 1, pp. 1–16, 27 Jan. 2016.

[53] M. Bertozzi and A. Broggi, "GOLD: a parallel real-time stereo vision system for generic obstacle and lane detection," *IEEE Trans. Image Process.*, vol. 7, no. 1, pp. 62–81, 1998.

[54] A. Broggi, P. Cerri, S. Ghidoni, P. Grisleri, and H. G. Jung, "Localization and analysis of critical areas in urban scenarios," in *Intelligent Vehicles Symposium, 2008 IEEE*. ieeexplore.ieee.org, Jun. 2008, pp. 1074–1079.

[55] C. Tu, B. J. van Wyk, Y. Hamam, K. Djouani, and S. Du, "Vehicle position monitoring using hough transform," *IERI Procedia*, 2013.

[56] V. Gaikwad and S. Lokhande, "Lane departure identification for advanced driver assistance," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 910–918, Apr. 2015.

[57] E. Koukoumidis, M. Martonosi, and L. S. Peh, "Leveraging smartphone cameras for collaborative road advisories," *IEEE Trans. Mob. Comput.*, vol. 11, no. 5, pp. 707–723, May 2012.

[58] J. Apple, P. Chang, A. Clauson, H. E. Dixon, and others, "Green driver: AI in a microcosm," in *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*. pdfs.semanticscholar.org, 2011.

[59] L. Andreone, A. Amditis, E. Deregibus, S. Damiani, D. Morreale, and F. Bellotti, "Beyond Context-Awareness: Driver-Vehicle-Environment adaptivity. from the COMUNICAR project to the AIDE concept," *IFAC Proceedings Volumes*, vol. 38, no. 1, pp. 109–114, 2005.

[60] J. Jo, S. J. Lee, H. G. Jung, K. R. Park, and J. Kim, "Vision-based method for detecting driver drowsiness and distraction in driver monitoring system," *Opt. Eng.*, vol. 50, no. 12, pp. 127 202–127 202–24, 1 Dec. 2011.

[61] B.-G. Lee and W.-Y. Chung, "A smartphone-based driver safety monitoring system using data fusion," *Sensors*, vol. 12, pp. 17 536–17 552, 17 Dec. 2012.

[62] D. F. Dinges and R. Grace, "Perclos: a valid psychophysiological measure of alertness as assessed by psychomotor vigilance," National HighwayTraffic Safety Administration , Office of Motor Carriers, Indianapolis, IN, Tech. Rep. Final Report 1998, Report No. FHWA-MCRT-98-006, 1998.

[63] S. Jerritta, M. Murugappan, R. Nagarajan, and K. Wan, "Physiological signals based human emotion recognition: a review," in *Signal Processing and its Applications (CSPA), 2011 IEEE 7th International Colloquium on.* ieeexplore.ieee.org, Mar. 2011, pp. 410–415.

[64] G. Li and W.-Y. Chung, "Detection of driver drowsiness using wavelet analysis of heart rate variability and a support vector machine classifier," *Sensors*, vol. 13, no. 12, pp. 16 494–16 511, 2 Dec. 2013.

[65] S. Ros-Aguilar, J. Merino, A. Snchez, and others, "Variation of the heartbeat and activity as an indicator of drowsiness at the wheel using a smartwatch," *International Journal of*, 2015.

[66] B. Warwick, N. Symons, X. Chen, and K. Xiong, "Detecting driver drowsiness using wireless wearables," in *Mobile Ad Hoc and Sensor Systems (MASS), 2015 IEEE 12th International Conference on.* ieeexplore.ieee.org, Oct. 2015, pp. 585–588.

[67] G. Castignani, T. Derrmann, R. Frank, and T. Engel, "Driver behavior profiling using smartphones: A Low-Cost platform for driver monitoring," *IEEE Intell. Transp. Syst. Mag.*, vol. 7, no. 1, pp. 91–102, 2015.

[68] B. G. Simons-Morton, M. C. Ouimet, Z. Zhang, S. E. Klauer, S. E. Lee, J. Wang, R. Chen, P. Albert, and T. A. Dingus, "The effect of passengers and risk-taking friends on risky driving and crashes/near crashes among novice teenagers," *J. Adolesc. Health*, vol. 49, no. 6, pp. 587–593, Dec. 2011.

[69] V. L. Neale, T. A. Dingus, S. G. Klauer, J. Sudweeks, and M. Goodman, "An overview of the 100-car naturalistic study and findings," *National Highway Traffic Safety Administration, Paper*, no. 05-0400, 2005.

[70] R. Eenink, Y. Barnard, M. Baumann, X. Augros, and F. Utesch, "UDRIVE: the european naturalistic driving study," *Proceedings of the TRA*, 2014.

[71] Y. Barnard, F. Utesch, N. Nes, R. Eenink, and M. Baumann, "The study design of UDRIVE: the naturalistic driving study across europe for cars, trucks and scooters," *Eur. Transp. Res. Rev.*, vol. 8, no. 2, pp. 1–10, 19 May 2016.

[72] C. Carney, D. McGehee, K. Harland, M. Weiss, and M. Raby, "Using naturalistic driving data to assess the prevalence of environmental factors and driver behaviors in teen driver crashes," 2015.

[73] G. M. Fitch, S. A. Soccolich, F. Guo, J. McClafferty, Y. Fang, R. L. Olson, M. A. Perez, R. J. Hanowski, J. M. Hankey, and T. A. Dingus, "The impact of hand-held and hands-free cell phone use on driving performance and safety-critical event risk," Tech. Rep., 2013.

[74] D. A. Norman and S. W. Draper, *User Centered System Design: New Perspectives on Human-Computer Interaction*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 1986.

[75] B. Downe-Wamboldt, “Content analysis: method, applications, and issues,” *Health Care Women Int.*, vol. 13, no. 3, pp. 313–321, Jul. 1992.

[76] J. Kawakita, “The original KJ method,” *Tokyo: Kawakita Research Institute*, 1991.

[77] M. Q. Patton, *Qualitative evaluation and research methods, 2nd ed.* Sage Publications, Inc, 1990.

[78] M. N. Marshall, “The key informant technique,” *Family Practice*, vol. 13, no. 1, pp. 92–97, Feb. 1996.

[79] J. Krajewski, D. Sommer, and others, “Steering wheel behavior based estimation of fatigue,” *of the Fifth...*, 2009.

[80] T. Chakravarty, A. Ghose, C. Bhaumik, and A. Chowdhury, “MobiDriveScore—A system for mobile sensor based driving analysis: A risk assessment model for improving one’s driving,” in *Sensing Technology (ICST), 2013 Seventh International Conference on*. ieeexplore.ieee.org, 2013, pp. 338–344.

[81] Y. Wang, J. Yang, H. Liu, Y. Chen, M. Gruteser, and R. P. Martin, “Sensing vehicle dynamics for determining driver phone use,” in *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys ’13. New York, NY, USA: ACM, 2013, pp. 41–54.

[82] M. Rezaei, M. Terauchi, and R. Klette, “Robust vehicle detection and distance estimation under challenging lighting conditions,” *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2723–2743, Oct. 2015.

[83] R. K. Satzoda, S. Lee, F. Lu, and M. M. Trivedi, “Snap-DAS: A vision-based driver assistance system on a snapdragon TM embedded platform,” in *Intelligent Vehicles Symposium (IV), 2015 IEEE*. ieeexplore.ieee.org, 2015, pp. 660–665.

[84] Shneiderman, B And Plaisant,, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Reading, MA: Addison-Wesley Publ. Co., 2010, vol. Fifth Edition.

[85] L. Stenneth, O. Wolfson, P. S. Yu, and B. Xu, “Transportation mode detection using mobile phones and GIS information,” in *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ser. GIS ’11. New York, NY, USA: ACM, 2011, pp. 54–63.

[86] W. Iso, “9241-11. ergonomic requirements for office work with visual display terminals (VDTs),” *The international organization for standardization*, vol. 45, 1998.

[87] J. Brooke and Others, “SUS-A quick and dirty usability scale,” *Usability evaluation in industry*, vol. 189, no. 194, pp. 4–7, 1996.

[88] T. S. Tullis and J. N. Stetson, “A comparison of questionnaires for assessing website usability,” in *Usability Professional Association Conference*. Citeseer, 2004, pp. 1–12.

[89] A. Bangor, P. Kortum, and J. Miller, “Determining what individual SUS scores mean: Adding an adjective rating scale,” *J. Usability Studies*, vol. 4, no. 3, pp. 114–123, May 2009.

[90] A. Bangor, P. T. Kortum, and J. T. Miller, “An empirical evaluation of the system usability scale,” *International Journal of Human–Computer Interaction*, vol. 24, no. 6, pp. 574–594, 2008.

[91] S. S. Shapiro and M. B. Wilk, “An analysis of variance test for normality (complete samples),” *Biometrika*, vol. 52, no. 3/4, pp. 591–611, 1965.

[92] W. Liu, J. C. Principe, and S. Haykin, *Kernel Adaptive Filtering: A Comprehensive Introduction*, ser. Adaptive and Cognitive Dynamic Systems: Signal Processing, Learning, Communications and Control. New York: Wiley, 2011.

[93] R. D. Foss and A. H. Goodwin, “Distracted driver behaviors and distracting conditions among adolescent drivers: findings from a naturalistic driving study,” *J. Adolesc. Health*, vol. 54, no. 5 Suppl, pp. S50–60, May 2014.

[94] ARMmbed, “Lis302 accelerometer,” https://developer.mbed.org/cookbook/LIS302-Accelerometer, 05.02.2017.

[95] intorobotics, “Most common and most budgeted arduino light sensors,” https://www.intorobotics.com/common-budgeted-arduino-light-sensors/, 05.02.2017.

[96] sainsmart, “Sainsmart mq135 sensor air quality,” http://www.sainsmart.com/sainsmart-mq135-sensor-air-quality-sensor-hazardous-gas-detection-module-arduino.html, 05.02.2017.

[97] amazon.fr, “Heart rate sensor for arduino,” https://www.amazon.fr/Capteur-fr%C3%A9quence-cardiaque-pulsesensor-Raspberry/dp/B01CBGH4N6, 05.02.2017.

[98] Forschungsinstitut für Kraftfahrwesen und Fahrzeugmotoren Stuttgart , “Stuttgart driving simulator,” http://www.fkfs.de/en/automotive-mechatronics/leistungen/driving-simulators/, 15.07.2016.

[99] Ergoneers, “Vehicle testing kit,” http://www.ergoneers.com/en/hardware/vehicle-testing-kit/, 04.08.2016.

[100] R. R. H. Farag, “Active Neuro-Fuzzy integrated vehicle dynamics controller to improve the vehicle handling and stability at complicated maneuvers,” Ph.D. dissertation.

[101] Automatic Labs, “Automatic lite,” https://www.automatic.com/, 03.08.2016.

[102] Wayray SA, “Wayray element,” https://wayray.com/element/, 03.08.2016.